

Descripteur de forme basé sur la théorie de Morse

par

Mehdi ZIANI

**mémoire présenté au Département d'informatique
en vue de l'obtention du grade de maître ès sciences (M.Sc.)**

**FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE**

Sherbrooke, Québec, Canada, Juillet 2011



**Library and Archives
Canada**

**Published Heritage
Branch**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque et
Archives Canada**

**Direction du
Patrimoine de l'édition**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence
ISBN: 978-0-494-83702-3

Our file Notre référence
ISBN: 978-0-494-83702-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Le 28 juillet 2011

*le jury a accepté le mémoire de Monsieur Mehdi Ziani
dans sa version finale.*

Membres du jury

Professeur Madjid Allili
Directeur de recherche
Département d'informatique

Professeur Djemel Ziou
Codirecteur de recherche
Département d'informatique

Professeur Layachi Bentabet
Évaluateur interne au programme
Département d'informatique
Université Bishop's

Professeur Shengrui Wang
Président rapporteur
Département d'informatique

Sommaire

Le principal objectif de ce mémoire de maîtrise est la conception et la validation d'un système de description et d'analyse de la forme, adapté pour tout ensemble de données multidimensionnel.

Après avoir étudié les différentes techniques utilisées précédemment dans ce domaine, nous avons démontré que la combinaison de la théorie de Morse avec la topologie algébrique et plus spécifiquement de l'homologie, était fortement pertinente pour construire un descripteur performant de formes pouvant être modélisées par des variétés de Morse.

Notre descripteur caractérise la forme pour toutes les paires (\mathcal{M}, φ) , où \mathcal{M} est une variété fermée et lisse et φ est la fonction de Morse définie sur \mathcal{M} . Cette description fait, ainsi, le lien entre l'évolution des points critiques de φ et les changements d'homologie relative sur la forme étudiée. Ces changements sont exprimés en terme de paires de niveaux de φ sur \mathcal{M} .

Lors de la validation de ce descripteur par la conception d'un système de recherche d'images par le contenu, plusieurs expériences ont été faites. De nouvelles mesures de distances ont été appliquées et une stratégie de combinaison des fonctions de Morse a donné des résultats expérimentaux très concluants.

Remerciements

Je tiens, en premier lieu, à remercier mon directeur le professeur Madjid Allili qui, tout au long de ma maîtrise, a su me donner le soutien moral et technique pour avancer, malgré les embûches qu'un étudiant peut rencontrer pendant un projet comme celui-là. Je lui suis très reconnaissant aussi pour la patience, le dévouement et la disponibilité qu'il a montré pour m'introduire et me familiariser avec des concepts et des théories très pointues qui m'étaient inconnues, notamment, le vaste domaine de la topologie computationnelle. Tant sur le plan académique qu'humain, nos échanges m'ont grandement enrichi et encouragé.

Je remercie également mon co-directeur le professeur Djemel Ziou pour m'avoir permis de poursuivre mes études dans des conditions favorables, notamment son support sur le plan matériel et de l'environnement de travail. Je le remercie aussi pour les concepts théoriques qu'il a su nous faire passer dans son cours.

Je n'oublie pas de remercier aussi David Corriveau. Son expertise et ses conseils m'ont été d'une précieuse aide.

De très chaleureux remerciements vont aussi aux membres du jury, les professeurs Shengrui Wang et Layachi Bentabet qui ont accepté de prendre de leur si précieux temps pour corriger ce mémoire.

Enfin, une pensée affective va à mes parents et à ma famille, qui m'ont toujours été d'un support essentiel.

Table des matières

Sommaire	ii
Remerciements.....	iii
Table des matières	iv
Liste des tableaux.....	vii
Liste des figures	viii
Introduction.....	1
Chapitre 1 Revue des descripteurs de forme	4
1.1 Les descripteurs de forme	4
1.2 Revue des types de descripteurs de formes :.....	5
1.2.1 Méthodes basées sur le contour	9
1.2.2 Méthodes basées sur la région	30
1.3 Mesures de similarité	39
1.4 Evaluation des résultats	45
1.4.1 La paire <i>Précision/Rappel</i>	45
1.4.2 Pourcentage de résultats pondérés (PWH).....	46
1.4.3 Pourcentage du classement selon la similarité (PSR)	47
1.4.4 Moyenne du rang de recherche normalisé	48
1.5 Conclusion.....	49
Chapitre 2 Topologie, homologie et théorie de Morse.....	50
2.1 Introduction	50
2.2 Topologie algébrique.....	51

2.2.1	Homéomorphismes et variétés	52
2.2.2	Equivalence homotopique.....	53
2.3	Homologie.....	55
2.4	Homologie cubique	56
2.5	Les groupes d'homologie	60
2.6	Calcul d'homologie	63
2.6.1	Méthode classique.....	63
2.6.2	Méthode de calcul par la réduction	64
2.7	Homologie relative	70
2.8	La suite exacte d'homologie d'une paire	73
2.9	La séquence exacte d'homologie d'un triplet	75
2.10	Théorie de Morse en imagerie	76
2.10.1	Points critiques et fonctions de Morse	76
2.10.2	Lemme de Morse	78
2.10.3	Attacher une k -cellule	81
2.10.4	Théorèmes de Morse.....	83
2.11	Conclusion.....	87
Chapitre 3 Descripteur de forme basé sur la théorie de Morse		88
3.1	Introduction	88
3.2	Descripteur de Morse	88
3.3	Propriétés et calculabilité du descripteur de Morse	92
3.3.1	Invariance.....	92
3.3.2	Relation avec des descripteurs topologiques connus.....	94
3.3.3	Calculabilité du MSD	96
3.4	Algorithme	100
3.5	Mesure de distance	100
3.6	Combinaison des fonctions de Morse	104
3.7	Présentation des résultats expérimentaux.....	108
3.7.1	Première expérience.....	110

3.7.2	Deuxième expérience.....	111
3.7.3	Troisième expérience	111
3.7.4	Quatrième expérience	113
3.8	Conclusion.....	115
Conclusion		116
Bibliographie.....		118

Liste des tableaux

Tableau 3.1 – Échantillons de chaque classe de notre base de données	109
Tableau 3.2 – Précisions moyennes avec et sans apprentissage	114

Liste des figures

Figure 1.1 - Cheminement d'un descripteur de forme	5
Figure 1.2 – Illustration de la région (b), les multi-contours (c), le contour externe (d) et la collection de points (e), de l'objet dans l'image originale (a).	8
Figure 1.3 - Classification des descripteurs de formes selon leur approche	9
Figure 1.4 - Ratio de circularité	12
Figure 1.5 – Centroïde (symbolisé par l'étoile) du contour. Les points du contour sont distribués de manière uniforme dans (a) et non-uniforme dans (b)	13
Figure 1.6 - <i>Bounding Box</i> minimale et les paramètres de l'élongation.	15
Figure 1.7 - Image Binaire (a)discrete et (b)continue.....	16
Figure 1.8 – Paramétrisation d'une droite avec son angle avec l'axe des x et sa distance de l'origine.....	17
Figure 1.9 – Nature des axes principaux.....	18
Figure 1.10 – Variance d'ellipse	19
Figure 1.11 – Une forme de pomme et sa signature de distance au centroïde	19
Figure 1.12 – Illustration de la représentation par <i>chain code</i>	24
Figure 1.13 – Illustration de la <i>chain code</i> différentielle (normalisée).....	25
Figure 1.14 – Exemple de décomposition polygonale.....	26
Figure 1.15 – Vecteurs chaîne: (a) description originale de la frontière; (b) Description normalisée	26
Figure 1.16 – (a) est l'orientation du jeton ; (b) Une forme de cheval décomposée en plusieurs jetons. Les nombres correspondants à chaque jeton sont la courbure et l'orientation de celui-ci.....	28

Figure 1.17 – Indexation des jetons dans un arbre <i>M-aire</i>	29
Figure 1.18 – Deux exemples d’image polaire et sa représentation cartésienne. (a) Image originale à coordonnées polaires ; (b) L’image polaire dans (a) représentée dans un espace cartésien.....	34
Figure 1.19 – (a) Echantillonnage en coordonnées polaires ; (b) Une étoile formée de barres (c) Forme rectangulaire.....	35
Figure 1.20 – Trois formes avec des nombres d’Euler différents.....	36
Figure 1.21 – (a) Enveloppe convexe et ses concavités ; (b) Représentation en arbre des concavités de l’enveloppe.....	38
Figure 1.22 – Construction de l’axe médian d’une forme rectangulaire	39
Figure 1.23 (a) La distance de Manhattan $d_{Man} = d_{11} + d_{12}$ et la distance à norme infinie $L_{\infty} = \max(d_{12}, d_{11}) = d_{12}$; (b) Distance Euclidienne $d_{Eu} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$; (c) les vecteurs A et A_I ont la même distance cosinus par rapport au vecteur B	42
Figure 1.24 – Distance de Hausdorff.....	44
Figure 2.1 – (a) une forme consistant en une seule composante connexe ; (b) forme divisée en deux composantes connexes	51
Figure 2.2 - L’homéomorphisme f transforme un voisinage de p sur le tore en un voisinage de l’origine du plan. Une telle transformée existe pour tout point du tore, sachant que le tore est une variété de dimension 2.....	53
Figure 2.3 - Le cercle est un retract de déformation de l’anneau	54
Figure 2.4 - Exemple de complexes, (a) bidimensionnel et (b) tridimensionnel. Dans (a) le complexe est formé par les carrés grisés.....	58
Figure 2.5 – Différents scénarios pour illustrer le calcul des groupes d’homologie	62
Figure 2.6	63
Figure 2.7 – Séquence d’effondrements élémentaires de $[0,1] \times [0,1] \subset \mathbb{R}^2$	66
Figure 2.8 – Algorithme HOMOLOGY	69
Figure 2.9 – Procédure REDUCE()	70
Figure 2.10 - L’intuition derrière l’homologie relative.....	72

Figure 2.11 - Les groupes d'homologie dans le complexe (a) modulo les sommets V_0 et V_1 sont isomorphes aux groupes d'homologie réduite de l'espace quotient dans (b) duquel on peut directement voir que $H_0(X, A) = 0$ et $H_1(X, A) = \mathbb{Z}$	73
Figure 2.12 - $\varphi(x) = x^3$ a un point critique dégénéré en 0.	77
Figure 2.13 - Exemples (a) : points critiques dégénérés, (b) : points critiques non- dégénérés	77
Figure 2.14 - Exemple de l'indice d'un point critique selon la dimension de sa variété instable	78
Figure 2.15 - Plonger un beigne dans une tasse de café. (b) dès que le beigne touche la surface du café, la portion de sa surface immergée change de l'ensemble vide à une forme homéomorphe au disque. (c) à l' instant où le café atteint le trou du beigne, la topologie de la portion immergée change d'un disque à un cylindre tronqué comme dans (d). À l' instant où le trou est immergé en entier comme dans (e), la topologie de la portion immergée change du cylindre tronqué à un tore tronqué comme dans (f). En relâchant le beigne, le café recouvrera tout le beigne, ce qui complètera la forme du tore	80
Figure 2.16 - les différentes classes d'homotopie du tore coupé	81
Figure 2.17 - L'attachement d'une 1-cellule à une 3-boule conduit à une forme homotopique au tore.....	82
Figure 2.18 - Exemple du théorème 2.4.....	84
Figure 2.19 – Exemple du théorème 2.5	84
Figure 2.20 - Exemple de la décomposition d'une variété en cellules élémentaires	85
Figure 2.21 - Exemple d'une λ -cellule élémentaire qui affecte l'homologie du niveau $\lambda - 1$	86
Figure 2.22 - Exemple d'une λ -cellule élémentaire qui affecte l'homologie du niveau λ ...	87
Figure 3.1 - Exemple d'une variété bidimensionnelle et la fonction hauteur discrétisée en neuf niveaux, ainsi que le descripteur de Morse associé défini par trois matrices	89
Figure 3.2 – Exemple d'une forme type contour et les matrices associées du descripteur de forme MSD	90

Figure 3.3 - dans (a) il est facile de voir que $MSD_f(x-\varepsilon, y+\varepsilon, 0)=1$ alors que $MSD_g(x, y, 0)=0$. Dans (b) $MSD_f(x-\varepsilon, y+\varepsilon, 0)=0$ alors que $MSD_g(x, y, 0)=1$. Mêmes observations à propos de $MSD_f(x+\varepsilon, y-\varepsilon, 0)$ et $MSD_g(x, y, 0)$.	97
Figure 3.4 – Exemple de la notion de persistance	102
Figure 3.5 – Exemple de la persistance versus contribution.	103
Figure 3.6 - Importance du choix adéquat de la fonction de mesure	105
Figure 3.7 - Aperçu de 5 résultats à deux requêtes dans notre outil	110
Figure 3.8 - Précision x Rappel : Plus haute est la courbe meilleur est le descripteur	112
Figure 3.9 - Précision x rappel pour les descripteurs décrit ci-dessus	113

Introduction

Les récentes avancées technologiques dans les domaines d'acquisition, de traitement et de stockage des données images, ont causé une formidable démultiplication du nombre et de la taille des bases de données d'images. La structure des systèmes d'analyse, de classification et de recherche d'images par le contenu, devient alors un véritable défi technique.

Dans ces systèmes de recherche d'images par le contenu (CBIR pour *Content-Based Image Retrieval*), les propriétés des images sont, généralement, décrites par la forme, la couleur ou la texture de l'objet/région dans l'image. Un système de CBIR est formé essentiellement d'une base de données et d'un descripteur de forme. Ce dernier consiste en la paire: vecteur de propriétés et distance métrique, qui sont utilisés pour indexer et classifier les images selon leurs similarités. Pour une requête spécifiée par l'utilisateur, le CBIR vise à retourner un groupe d'images similaires à la requête selon les critères choisis. Lors de la recherche, on désire optimiser tant l'efficacité (qualité de la réponse) que la performance (complexité et temps de calcul) du système. Clairement, l'efficacité et la performance d'un système d'analyse et de classification d'image vont dépendre du choix du descripteur de forme. Ce choix est généralement motivé par les performances visées, mais aussi par la nature et le type d'images dont on dispose.

Les propriétés extraites d'une image peuvent aller de critères simples et de bas niveau tels que le périmètre, la circularité ou les moments statistiques, jusqu'à des propriétés beaucoup plus complexes, telles que la décomposition selon la courbure [60] où les propriétés topologiques de la forme telles que l'homologie. L'attention est portée dans ce mémoire sur les descripteurs à base de propriétés topologiques de la forme. Celles-ci sont connues pour être moins sensibles aux variations qui peuvent subvenir au niveau de l'apparence de la

forme. Ces variations peuvent être dues au bruit, mais aussi à diverses déformations. Outre cette grande résistance aux distorsions, les descripteurs topologiques permettent aussi de réduire la quantité de données utilisées pour décrire une forme tout en en gardant assez pour une très bonne caractérisation.

On peut citer comme descripteurs topologiques, celui à base de la *fonction de taille* [64], le *graphe d'évolution d'homologie* [65] et le *descripteur de forme basé sur la théorie de Morse*[53]. Et c'est effectivement ce dernier que nous nous sommes proposé d'en valider l'approche et la pertinence, en l'utilisant dans un système de recherche d'images par le contenu. Nous souhaitons, en effet, à travers ce mémoire, faire connaître la puissance de la topologie computationnelle et contribuer à la compréhension des mécanismes de celle-ci, lorsqu'elle est appliquée au domaine de l'imagerie.

Dans le premier chapitre de notre mémoire, nous voulions, au début, faire une brève catégorisation des descripteurs de forme existants. Mais nous nous sommes vite rendu compte que le type de catégorisation, lui-même, était sujet à débat. En effet, il existe dans la littérature une multitude de caractéristiques qui peuvent influencer le moyen de classifier un descripteur. Nous avons donc décidé de présenter ces différentes classifications et nous expliquons pourquoi nous pensons que celle décrite dans [15] était, selon nous, la plus judicieuse. Nous nous attelons ensuite, à présenter dans le détail et pour chaque classe, certains des descripteurs les plus connus.

Nous nous proposons dans le deuxième chapitre de présenter les notions de base les plus utiles à la compréhension de la topologie algébrique, l'homotopie, l'homologie, l'homologie relative et bien sur la théorie de Morse. Plusieurs propositions et théorèmes y sont présentés pour arriver à faire le lien entre l'évolution de l'apparition de points critiques d'une fonction de Morse définie sur une forme et les changements de topologie de celle-ci, ce qui est la base même du descripteur de forme basé sur la théorie de Morse. Seront présentés aussi, certains résultats, très puissant, qui lient les points critiques de la fonction de Morse aux générateurs d'homologie réduite d'une série de niveaux de la forme étudiée.

Dans le dernier chapitre nous commençons par définir formellement *le descripteur de forme basé sur la théorie de Morse*. On présente par la suite ses propriétés qui font qu'il caractérise les points critiques d'une fonction. Cette caractérisation se fait par le calcul des générateurs d'homologie relative d'une paire de niveaux sur la forme. Ceci sera, ensuite, sauvegardé sous forme de matrices correspondant aux différents niveaux d'homologie. Ces matrices formeront, ainsi, la description de la forme. Nous proposons dans ce chapitre une mesure de distance qui permet de palier à l'apparition de valeurs aberrantes dans les matrices. Nous y abordons la difficulté de choisir au préalable, quelle fonction de Morse à utiliser pour quel type d'image. Et nous expliquons la méthode heuristique adoptée pour combiner plusieurs fonctions de Morse lors de comparaisons d'images.

Toujours dans le même chapitre, nous effectuons plusieurs expérimentations sur notre système de recherche d'images par le contenu. Ces expérimentations, même comparées à des standards du domaine tels que le CSS de MPEG-7 [60], donnent des résultats très satisfaisants et prouvent la pertinence de l'utilisation des descripteurs topologiques dans le domaine de l'analyse, la recherche et la reconnaissance de forme.

Enfin, en guise de conclusion, nous faisons un résumé de ce qui a été effectué et nous présentons des pistes à explorer pour les futurs travaux dans le domaine.

Chapitre 1

Revue des descripteurs de forme

Ce chapitre sera dédié à la présentation des différentes approches utilisées dans la construction d'un descripteur de forme. Mais tout d'abord **qu'est-ce qu'un descripteur de forme?**

1.1 Les descripteurs de forme

Un descripteur est un ensemble de nombres organisés de diverses manières : valeurs simples, vecteurs, matrices ou même des graphes [1]. Ces nombres sont produits pour décrire une forme donnée, cette dernière peut ne pas être entièrement reconstituée à partir de ce descripteur mais les descripteurs de formes différentes devront être suffisamment différents pour pouvoir dissocier ces formes. Un descripteur de forme suit le cheminement schématisé dans la Figure 1.1

Qu'est-ce qu'un bon descripteur de forme?

Plus la différence entre deux descriptions de deux formes différentes, est significative et moins elle l'est pour deux formes similaires, meilleur est le descripteur.

Qu'est-ce que la similarité entre deux formes ?

En fait personne n'a pu jusqu'à présent répondre à cette question de manière assez convaincante. Ainsi, si nous pouvions quantifier la similarité entre formes, nous aurions le descripteur parfait; En réalité ce que fait le descripteur ce n'est qu'essayer de quantifier la

forme d'une manière qui concorde avec la perception et l'intuition humaine (qui est fortement subjective).

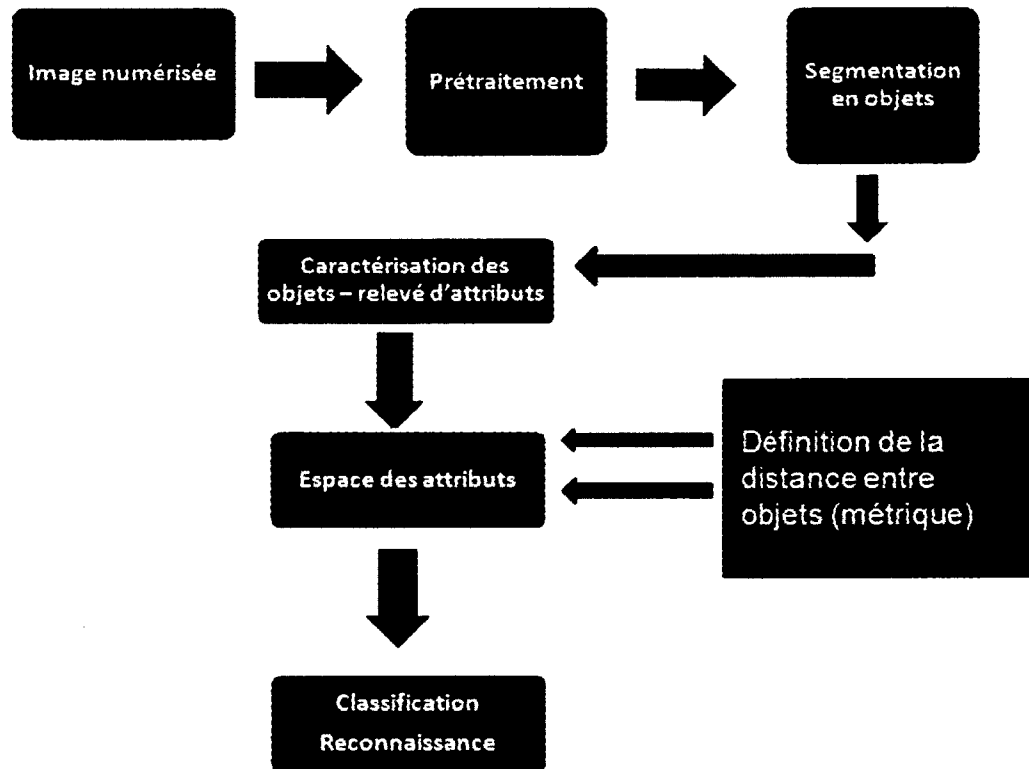


Figure 1.1 - Cheminement d'un descripteur de forme

1.2 Revue des types de descripteurs de formes :

Les méthodes d'analyse et de description de forme peuvent être classifiées selon plusieurs critères [2]. Pavlidis [3] a proposé les classifications suivantes :

La première de ces classifications est basée sur l'utilisation des points de contours versus ceux de l'intérieur de la forme pour décrire la forme. Les deux classes d'algorithmes résultantes sont connues comme méthodes externes (basées sur les contours) et les méthodes internes (ou globales).

Des exemples d'algorithmes qui décomposent les contours de la forme peuvent être trouvés dans [4, 5, 6], d'autres utilisant la transformée de Fourier du contour dans [7, 8].

Des exemples de méthodes globales, telle que celles utilisant la transformée de l'axe médian (appelé aussi axe de symétrie) proposée par Blum et appelée méthodes *MAT* (pour *Medial Axis Transform*) sont décrites dans [9, 4, 5] . D'autres dans cette même catégorie utilisent les moments statistiques [10, 11], ou encore, celles basées sur la décomposition de la forme en des formes plus basiques [12, 13].

Une autre classification des algorithmes de description de forme peut être faite sur la base de la nature du résultat produit, qu'il soit numérique ou non-numérique. La *MAT*, par exemple, produit une autre image (contenant l'axe de symétrie). D'un autre côté, les transformées numériques produisent des nombres (scalaires ou vecteurs) comme résultat. Des exemples de ceux-ci sont notamment des transformées de Fourier [7, 8] et celles basées sur les moments statistiques [10, 11].

Toujours selon Pavlidis, la troisième classification peut elle être faite sur une base de préservation (ou conservation) d'information. Ainsi on aura des descripteurs donnant une information assez précise pour permettre une reconstruction de la forme, appelés descripteurs préservateurs. Et d'autres descripteurs ne permettant qu'une reconstruction partielle, ces méthodes sont appelées méthodes non-préservatrices. Comme exemple d'information non-préservatrice on donne souvent le ratio Surface/Périmètre, ainsi plusieurs formes significativement différentes, peuvent avoir le même ratio mais la reconstruction de la forme originale à partir de ce ratio reste impossible.

Mingqiang [14] quant à lui prend en compte pour classifier les descripteurs de forme, la définition de l'apparence d'un objet physique. Pour lui, celle-ci est la silhouette ou la configuration de la surface caractérisant l'objet physique. Cette configuration caractéristique peut être représentée par son contour, sa région ou encore par un ensemble fini de points.

- Contour : la frontière sera ainsi exploitée pour représenter l'objet par une courbe continue et fermée dans le plan Figure 1.2(c) (multi-contours) et Figure 1.2(d)

(contour externe). On pourra, dès lors, considérer l'objet comme étant une collection de tous les points frontière de la région et défini par l'équation paramétrique discrète, dans un espace de coordonnées cartésiennes (x,y) , suivante :

$$\Gamma(n) = (x(n), y(n))$$

où $n \in [0, N-1]$; un contour pourra être paramétrisé par n'importe quel nombre N de sommets avec $\Gamma(N) = \Gamma(0)$.

- Région : formée de tous les pixels à l'intérieur de la forme, elle peut être constituée d'une seule ou d'un ensemble de régions et de trous de l'objet Figure 1.2(b). En imagerie et plus particulièrement dans le domaine de la reconnaissance de forme, la région est décrite par la fonction suivante : $f(x,y) = \begin{cases} 1 & \text{si } (x,y) \in D \\ 0 & \text{Sinon} \end{cases} \quad (1.1)$

où D est le domaine de l'image binaire. En d'autres termes, la région est une image binaire où les points de l'objet sont à 1 et les points de l'arrière plan à 0.

- Un ensemble fini de points : ce sont des points sur le contour de l'objet qui sont pris pour obtenir une représentation de la forme Figure 1.2(e). Ils peuvent être des points séparés ou des ensembles de segments de lignes qui portent en eux une importance significative dans l'apparence de l'objet pour la vision humaine.

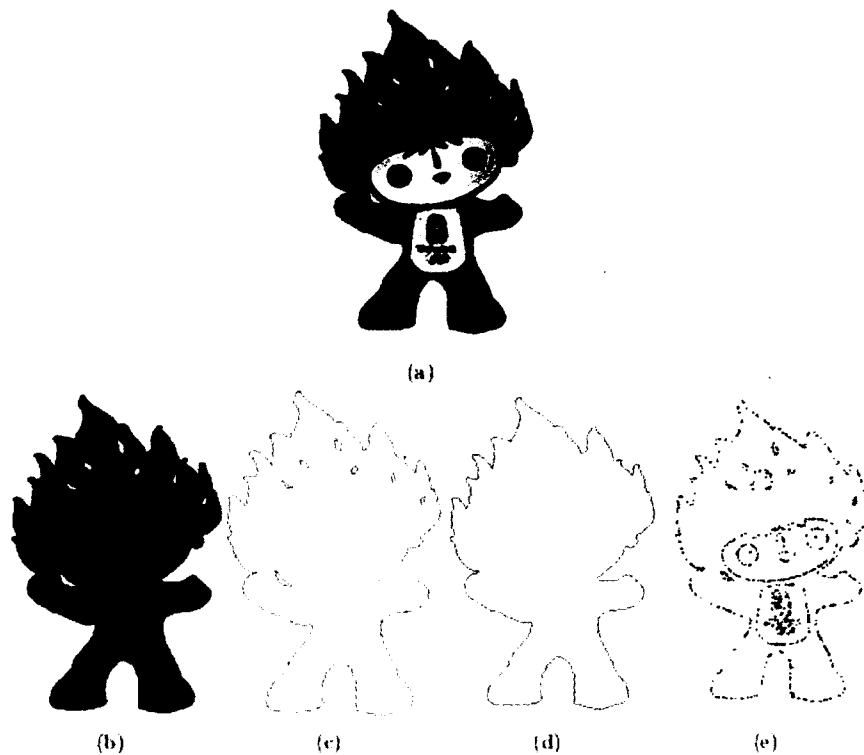


Figure 1.2 – Illustration de la région (b), les multi-contours (c), le contour externe (d) et la collection de points (e), de l'objet dans l'image originale (a).

Ainsi et selon ce découpage des différentes représentations d'un objet, on aura trois types de descripteurs basés sur le contour, sur les régions ou sur un ensemble de points choisis sur le contour.

Mais la classification des descripteurs la plus précise et pertinente selon nous, est celle proposée par Zhang et Lu dans [15], où ils basent leur découpage sur le fait que les attributs sont soit extraits du contour uniquement ou alors de l'entièreté de l'objet, donnant deux grandes classes de descripteurs de forme basés sur le contour ou sur les régions. Les auteurs proposent aussi de découper chacune de ces deux grandes classes en deux sous-classes qui décrivent deux approches différentes sous chacune de ces classes. Une structurelle et l'autre globale. Cette sous-division est basée sur le fait que l'objet sera représenté soit comme un

tout ou alors par des segments/sections (primitives). La hiérarchie complète est résumée dans la Figure 1.3.

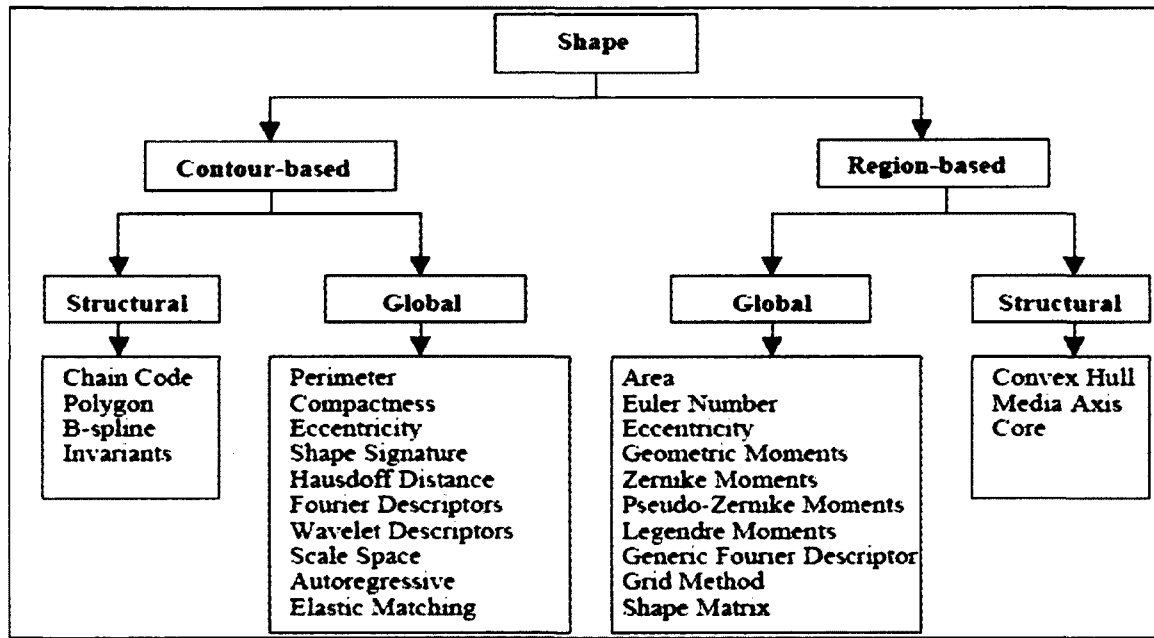


Figure 1.3 - Classification des descripteurs de formes selon leur approche

1.2.1 Méthodes basées sur le contour

Ces méthodes exploitent la frontière de l'objet. Elles sont généralement divisées en deux types d'approches très différentes dans la modélisation du contour ; une approche continue (globale) et une autre discrète (structurelle). La première ne divise pas la forme en sous-parties mais utilise plutôt un vecteur de propriétés dérivé de l'intégralité du contour pour décrire la forme. La mesure de similarité entre formes dans ce cas est généralement une distance métrique entre vecteurs. La deuxième approche (discrète) divise, quant à elle, le contour en segments appelés aussi primitives, selon un certain critère et la représentation se fait habituellement avec des chaînes ou des graphes (arbres), la mesure de similarité s'effectue donc avec des processus de mise en correspondance de chaînes ou de graphes.

Dans la suite nous présentons plus en détail des exemples de ces deux approches :

1.2.1.1 Méthodes globales

Ces méthodes calculent un vecteur d'attributs multidimensionnel à partir du contour de la forme. La mise en correspondance entre formes devient dès lors un simple processus de calcul de distances métriques telles que la distance euclidienne ou la *City Block* pour ne citer qu'elles. Une correspondance « point à point » peut aussi être envisagée. Voici à présent des exemples de ce type de descripteurs de forme :

i) **Descripteurs de forme basiques** : certaines propriétés géométriques peuvent être utilisées comme descripteurs de forme basiques [16]. Les descripteurs simples ne discriminent généralement que les formes ayant une très large différence mais sont cependant utilisés comme filtres lors de prétraitements pour éliminer les formes trop divergentes de la requête.. on peut citer parmi ceux-ci :

- **Périmètre** : Défini comme le nombre de pixels sur le contour de la forme
- **Moyenne de l'énergie de flexion** : elle est définie par EF :

$$EF = \frac{1}{N} \sum_{s=0}^{N-1} K(i)^2$$

où $K(i)$ est la fonction de courbure, s est le paramètre de longueur d'arc et N est le nombre de points du contour [2]. On peut facilement prouver que le cercle est la forme avec une moyenne d'énergie de flexion minimale.

- **Ratio de circularité** : Représente jusqu'à quel degré la forme se rapproche d'un cercle. Il en existe trois définitions :

1. C'est le ratio de la surface de la forme à la surface du cercle ayant le

même périmètre que la forme $C_1 = \frac{S_f}{S_c}$ où S_f est la surface de la

forme, S_c est la surface du cercle ayant le même périmètre que la

forme. En posant P comme périmètre, S_c sera donc $S_c = P^2 / 4\pi$, C_1 sera ainsi $C_1 = 4\pi \cdot S_f / P^2$ et vu que 4π est une constante, on aura la deuxième définition du ratio.

2. C'est le ratio de la surface de la forme au carré du périmètre de celle-ci

$$C_2 = \frac{S_f}{P^2}$$

3. Le ratio de circularité est appelé aussi variance de cercle et est défini

par :

$$C_{va} = \frac{\sigma_R}{\mu_R} \quad (1.2)$$

où μ_R et σ_R sont respectivement la moyenne et l'écart type de la distance radiale entre le centroid de la forme (g_x, g_y) et tous les points du contour (x_i, y_i) , $i \in [0, N-1]$, leurs formules respectives sont comme suit :

$$\mu_R = \frac{1}{N} \sum_{i=0}^{N-1} d_i \quad \text{et} \quad \sigma_R = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (d_i - \mu_R)^2}$$

Où $d_i = \sqrt{(x_i - g_x)^2 + (y_i - g_y)^2}$

La forme la plus compacte étant le cercle comme montré dans la Figure 1.4.

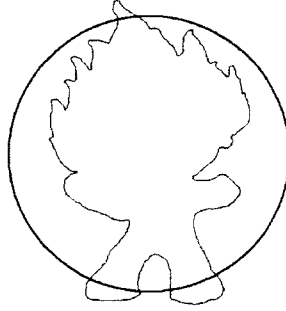


Figure 1.4 - Ratio de circularité

- **Centre de gravité :** Aussi appelé *Centroïde*, sa position est fixée dépendamment de la manière avec laquelle la forme est représentée. Si celle-ci est représentée par sa région alors comme dans l'équation (1), le centroïde sera donné par :

$$\begin{cases} g_x = \frac{1}{N} \sum_{i=0}^{N-1} x_i \\ g_y = \frac{1}{N} \sum_{i=0}^{N-1} y_i \end{cases}$$

où N est le nombre de points dans la forme $(x_i, y_i) \in \{(x_i, y_i) | f(x_i, y_i) = 1\}$.

Si la forme, par contre, est représentée par son contour, alors la position du centroïde sera donnée par :

$$\begin{cases} g_x = \frac{1}{6A} \sum_{i=0}^{N-1} (x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \\ g_y = \frac{1}{6A} \sum_{i=0}^{N-1} (y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \end{cases}$$

où A est la surface du contour donnée par :

$$A = \frac{1}{2} \left| \sum_{i=0}^{N-1} (x_i y_{i+1} - x_{i+1} y_i) \right|$$

La position du centroïde sera fixée avec différentes distributions des points sur le contour dans la Figure 1.5, on remarque que sa position reste la même malgré une distribution de points sur le contour différente.

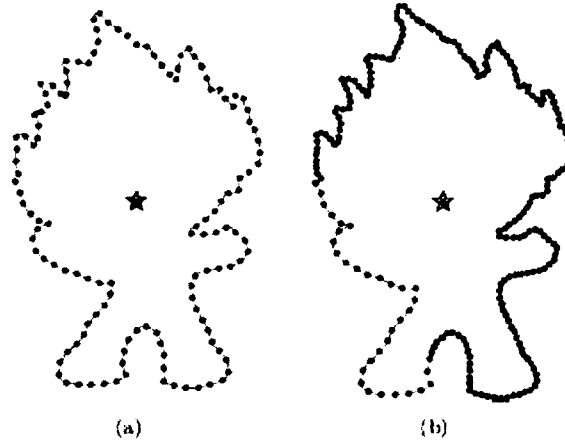


Figure 1.5 – Centroïde (symbolisé par l'étoile) du contour. Les points du contour sont distribués de manière uniforme dans (a) et non-uniforme dans (b)

- **Axe de moindre inertie** : Unique pour une forme, il sert comme référence pour décrire l'orientation de celle-ci. Le *ALI* (*Axis of least inertia*), est défini comme étant la ligne pour laquelle l'intégrale des carrés des distances aux points du contour, est minimale. Vu que cet axe passe par le centroïde du contour de la forme, pour trouver le *ALI* on fait habituellement une translation de manière à faire coïncider le centroïde avec l'origine d'un système de coordonnées cartésiennes. Considérons $x \sin \theta - y \cos \theta = 0$ comme l'équation paramétrique du *ALI*, l'angle d'inclinaison θ est estimé comme suit :

Soit α l'angle formé par le *ALI* et l'axe des x , l'inertie sera donnée par :

$$I = \frac{1}{2}(a+c) - \frac{1}{2}(a-c)\cos(2\alpha) - \frac{1}{2}b\sin(2\alpha)$$

$$\text{Où } a = \sum_{i=0}^{N-1} x_i^2, b = 2 \sum_{i=0}^{N-1} x_i y_i, c = \sum_{i=0}^{N-1} y_i^2$$

Et donc :

$$\frac{dI}{d\alpha} = (a-c)\sin(2\alpha) - b\cos(2\alpha)$$

$$\frac{d^2I}{d\alpha^2} = 2(a-c)\cos(2\alpha) + 2b\sin(2\alpha)$$

Mettons $dl/d\alpha = 0$ on obtient : $\alpha = \frac{1}{2} \arctan\left(\frac{b}{a-c}\right), -\frac{\pi}{2} < \alpha < \frac{\pi}{2}$

L'angle d'inclinaison θ sera donnée par :

$$\theta = \begin{cases} \alpha + \frac{\pi}{2} & \text{si } \frac{d^2 I}{d\alpha^2} \\ \alpha & \text{Sinon} \end{cases}$$

- **Excentricité** : C'est le ratio de la longueur de l'axe majeur à celle de l'axe mineur de la forme. Elle peut être calculée par la méthode des axes principaux ou par la méthode de la « *Bounding Box* » minimale.

1. Méthode des axes principaux : Les axes principaux d'une forme donnée sont les deux segments de droites qui s'intersectent de manière orthogonale au centroïde. Ils représentent les directions majeures avec une corrélation croisée [17]. De cette façon le contour est vu comme une instance d'une distribution statistique. Considérons la matrice de covariance C d'un contour :

$$C = \frac{1}{N} \sum_{i=0}^{N-1} \begin{pmatrix} x_i - g_x \\ y_i - g_y \end{pmatrix} \begin{pmatrix} x_i - g_x \\ y_i - g_y \end{pmatrix}^T = \begin{pmatrix} c_{xx} & c_{xy} \\ c_{yx} & c_{yy} \end{pmatrix} \quad (1.3)$$

Où :

$$\begin{aligned} c_{xx} &= \frac{1}{N} \sum_{i=0}^{N-1} (x_i - g_x)^2 \\ c_{xy} &= \frac{1}{N} \sum_{i=0}^{N-1} (x_i - g_x)(y_i - g_y) \\ c_{yx} &= \frac{1}{N} \sum_{i=0}^{N-1} (y_i - g_y)(x_i - g_x) \\ c_{yy} &= \frac{1}{N} \sum_{i=0}^{N-1} (y_i - g_y)^2 \end{aligned}$$

$G(g_x, g_y)$ est le centroïde de la forme clairement ici : $c_{xy} = c_{yx}$. Les longueurs des deux axes majeurs sont égales aux valeurs propres λ_1 et λ_2 de la matrice de covariance C respectivement. Le calcul de ces valeurs propres se fait de la manière suivante :

$$\det(C - \lambda_{1,2}I) = \det \begin{pmatrix} c_{xx} - \lambda_{1,2} & c_{xy} \\ c_{yx} & c_{yy} - \lambda_{1,2} \end{pmatrix} = (c_{xx} - \lambda_{1,2})(c_{yy} - \lambda_{1,2}) - c_{xy}^2 = 0$$

Et donc :

$$\begin{cases} \lambda_1 = \frac{1}{2} \left[c_{xx} + c_{yy} - \sqrt{(c_{xx} + c_{yy})^2 - 4(c_{xx}c_{yy} - c_{xy}^2)} \right] \\ \lambda_2 = \frac{1}{2} \left[c_{xx} + c_{yy} + \sqrt{(c_{xx} + c_{yy})^2 - 4(c_{xx}c_{yy} - c_{xy}^2)} \right] \end{cases}$$

L'excentricité sera finalement : $E = \lambda_2 / \lambda_1$

2. Méthode de la « *Bounding Box* » minimale : la *Bounding Box* est le plus petit rectangle qui peut contenir la forme. Pour une forme arbitraire, l'excentricité est le ratio entre la longueur L et la largeur W de la *Bounding Box* (voir Figure 1.6). Un autre concept dérivé de l'excentricité se nomme l'élongation. Elo est donnée par : $Elo = 1 - W/L$.

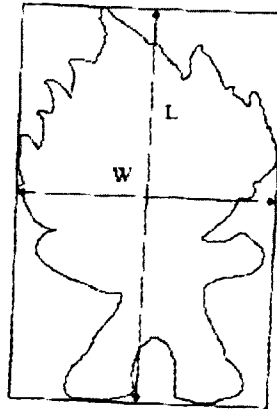


Figure 1.6 - *Bounding Box* minimale et les paramètres de l'élongation.

L'élongation prend ses valeurs dans $[0,1]$. Une forme symétrique par rapport à tous ses axes comme le cercle ou le carré par exemple, a une élongation nulle (égale à 0). Tandis qu'une forme qui a un large ratio d'aspect, cette forme aura une élongation plus proche de 1.

Relation entre centroïde, ALI et axes principaux : Nous présentons dans ce paragraphe le lien entre ces différents attributs d'un objet dans une image. La fonction caractéristique d'un objet dans une image étant :

$$f(x, y) = \begin{cases} 1 & \text{pour les points de l'objet} \\ 0 & \text{pour les points d'arrière plan} \end{cases}$$

La surface (l'aire) A de l'objet sera donnée par le moment d'ordre 0 de l'objet : Dans le cas où l'image (Figure 1.7) est continue (résolution infinie) $A = \iint f(x, y) dx dy$. Dans le cas discret : $A = \sum \sum b_{ij}$

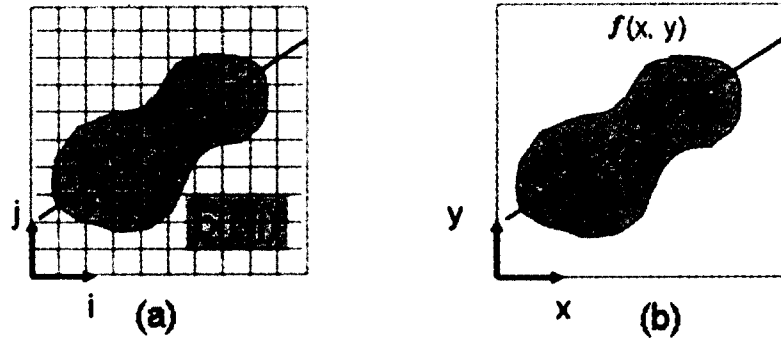


Figure 1.7 - Image Binaire (a)discrete et (b)continue

Le centre de masse (centroïde) sera donné par le moment d'ordre 1. En continu sa formule est $\bar{x} = \frac{1}{A} \iint x f(x, y) dx dy$ et $\bar{y} = \frac{1}{A} \iint y f(x, y) dx dy$; en discret cela donne $\bar{x} = \frac{1}{A} \sum \sum i b_{ij}$ et $\bar{y} = \frac{1}{A} \sum \sum j b_{ij}$.

L'axe de moindre inertie identifie l'orientation dominante de l'objet, il constitue en plus de ça un axe minimisant le moment de second ordre. Il minimise alors la fonction. $E = \iint r^2 f(x, y) dx dy$.

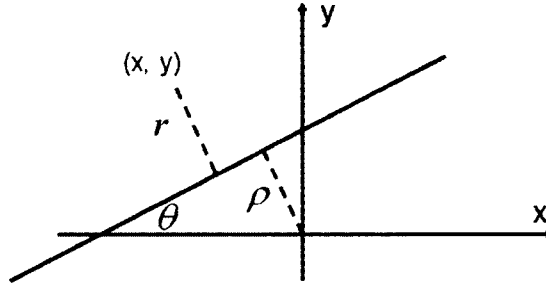


Figure 1.8 – Paramétrisation d'une droite avec son angle avec l'axe des x et sa distance de l'origine

L'équation paramétrique de cette droite peut être donnée sous la forme $x \sin \theta - y \cos \theta + \rho = 0$.

En résolvant ce problème de minimisation on obtient :

$$E = a \sin^2 \theta - b \sin \theta \cos \theta + c \cos^2 \theta \text{ où}$$

$$a = \iint (x')^2 f(x', y') dx' dy'$$

$$b = 2 \iint (x' y') f(x', y') dx' dy'$$

$$c = \iint (y')^2 f(x', y') dx' dy'$$

$$x' = x - \bar{x} \text{ et } y' = y - \bar{y}$$

a , b et c sont les moments de second degré. Ils paramètrent l'axe de moindre inertie, on remarque aussi qu'ils constituent les vecteurs de la fameuse matrice de

covariance $C = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$ dont les valeurs propres sont utilisées pour calculer

les axes majeurs. On obtient aussi lors de la résolution du problème de minimisation, deux solutions possibles pour θ l'une positive et l'autre négative qui vont définir la nature des axes de la manière décrite dans la Figure 1.9. L'axe

majeur sera de longueur λ_1 et son angle sera θ_1 . L'axe mineur sera de longueur λ_2 son angle sera θ_2 .

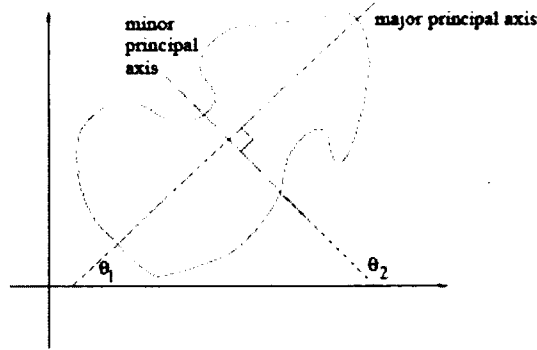


Figure 1.9 – Nature des axes principaux

N.B : dans le cas discret : $a = \sum \sum i^2 b_{ij}$, $b = 2 \sum \sum ij b_{ij}$ et $c = \sum \sum j^2 b_{ij}$.

- **Variance d'ellipse :** la variance d'ellipse E_{va} est le degré avec lequel notre forme épouse l'ellipse qui l'entoure Figure 1.10 et qui a la même matrice de covariance que celle de notre forme : $C_{ellipse} = C$ (cf Eq 1.3). En pratique l'approche inverse donne un rendement meilleur. Supposons que :

$$V_i = \begin{pmatrix} x_i - g_x \\ y_i - g_y \end{pmatrix}$$

$$d'_i = \sqrt{V_i^T \cdot C_{ellipse}^{-1} \cdot V_i}$$

$$\mu'_R = \frac{1}{N} \sum_{i=1}^{N-1} d'_i \quad \text{et} \quad \sigma'_R = \sqrt{\frac{1}{N} \sum_{i=1}^{N-1} (d'_i - \mu'_R)^2}$$

et donc

$$E_{va} = \frac{\sigma'_R}{\mu'_R}$$

En comparaison avec l'équation (1.2) (variance de cercle), E_{va} représente la forme de manière plus précise Figure 1.10.



Figure 1.10 – Variance d'ellipse

- ii) **Signature de la forme :** Une signature représente la forme par une fonction unidimensionnelle dérivée du contour de celle-ci. Plusieurs signatures existent, on peut citer par exemple la distance au centroïde Figure 1.11, l'angle tangent, le profil centroïdal, les coordonnées complexes, courbure, longueur de corde...etc.

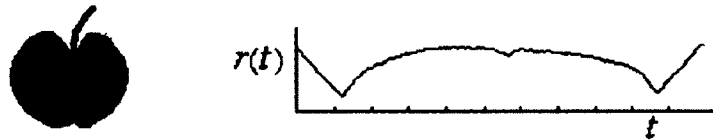


Figure 1.11 – Une forme de pomme et sa signature de distance au centroïde

Les signatures de forme sont généralement normalisées pour être invariantes à la translation et au changement d'échelle. Pour compenser les changements d'orientations, un *matching* des différents déplacements sera nécessaire pour trouver les meilleures correspondances entre formes[15,18]. Tous ces processus de normalisation rendent le matching trop coûteux pour une reconnaissance de forme On-Line. En plus de ce coût élevé, les signatures sont sensibles au bruit et des altérations sur le contour peuvent causer de larges erreurs dans le matching. Il est, dès lors, déconseillé de décrire directement et uniquement une forme par sa signature.

iii) **Les moments frontière** : Ils peuvent être utilisés pour réduire les dimensions de la représentation du contour. En supposant que celui-ci est représenté comme une signature de forme $z(i)$, le $r^{\text{ième}}$ moment m_r et le moment central μ_r sont estimés [14] comme ceci :

$$m_r = \frac{1}{N} \sum_{i=0}^{N-1} [z(i)]^r \quad \text{et} \quad \mu_r = \frac{1}{N} \sum_{i=0}^{N-1} [z(i) - m_1]^r$$

Où N est le nombre de points sur le contour. Les moments normalisés $\bar{m}_r = m_r / (\mu_2)^{r/2}$ et $\bar{\mu}_r = \mu_r / (\mu_2)^{r/2}$ sont invariants à la translation, la rotation et le changement d'échelle.

Des descripteurs moins sensibles au bruit peuvent aussi être obtenus par : $F_1 = (\mu_2)^{1/2} / m_1$, $F_2 = \mu_3 / (\mu_2)^{3/2}$ et $F_3 = \mu_4 / (\mu_2)^2$. L'avantage des moments frontière c'est leur facilité d'implémentation, mais par contre, il est très compliqué de donner des interprétations physiques aux moments de plus grands ordres.

iv) **Méthodes stochastiques** : Les modèles de séries temporelles et plus spécifiquement la modélisation autorégressive (AR), ont été utilisés pour calculer des descripteurs de forme [19,20]. Les méthodes dans cette classe sont basées sur une modélisation stochastique d'une fonction 1-D f , obtenue comme décrit dans la section « signature de forme ». Un modèle linéaire autorégressif exprime la valeur d'une fonction par une combinaison linéaire d'un certain nombre de valeurs précédentes. En effet, chaque valeur de la fonction dans la séquence a certaines corrélations avec des valeurs précédemment calculées et peut donc être prédite par disons M de ces observations précédentes. Le model autorégressif est donc un prédicteur simple de l'état actuel de la fonction par une combinaison linéaire des états précédents plus une constante et une erreur :

$$\hat{f}_t = \alpha + \sum_{j=1}^m \theta_j f_{t-j} + \sqrt{\beta} \omega_t$$

où les θ_j sont les coefficients du modèle AR, m est l'ordre du modèle, i.e., décrit combien de valeurs précédentes sont utilisées par le modèle. $\sqrt{\beta} \omega_t$ est l'erreur courante ou résiduelle décrivant la précision des prédictions. α est proportionnelle à la moyenne

des valeurs de la fonction. Les paramètres $\{\alpha, \theta_1, \dots, \theta_m, \beta\}$ sont estimés en utilisant le critère du moindre carré. Les θ_j estimés sont invariants à la translation, la rotation et au changement d'échelle. α et β ne sont, par contre, pas invariants à ce dernier, mais le quotient $\alpha/\sqrt{\beta}$ qui reflète le ratio du signal au bruit du contour est considéré comme invariant au changement d'échelle d'où l'utilisation du vecteur de propriétés $[\theta_1, \dots, \theta_m, \alpha/\sqrt{\beta}]^T$ comme descripteur de forme. L'inconvénient de la méthode *AR* est que dans le cas d'information très complexe, un petit nombre de paramètres *AR* ne sera pas suffisant pour une description adéquate de la forme. Le choix de m s'avère aussi compliqué, en plus du fait que la signification physique associée à chaque θ_j n'est pas claire.

- v) **Transformée spectrale** : Les descripteurs spectraux contournent le problème du bruit et des variations sur le contour en analysant la forme dans le domaine spectral. Ces descripteurs incluent le descripteur de Fourier (FD) et le descripteur à ondelettes (WD). Ils sont dérivés de transformées spectrales sur des fonctions 1-D décrivant des signatures de forme telles que décrites dans *ii*). Le descripteur de Fourier est encore utilisé bien que cette méthode soit vieille de 40 ans. Ceci est dû au fait qu'aussi bien la classification que la reconnaissance de forme usant du FD sont faciles à calculer et robustes au bruit.

En général le FD à une dimension est obtenu en appliquant la transformée de Fourier sur une signature de forme. Les coefficients normalisés de la T.F formeront le descripteur lors de la reconnaissance de forme. Il faut dire aussi que des FD dérivés de signatures de forme différentes présentent des performances inégales, ainsi et comme montré dans [21], le FD dérivé de la fonction de distance au centroïde $r(t)$ donne clairement de meilleures performances que les FD dérivées d'autres signatures de forme. La T.F discrète de $r(t)$ est donnée par :

$$a_n = \frac{1}{N} \sum_{t=0}^{N-1} r(t) \exp\left(\frac{-j2\pi nt}{N}\right), \quad n = 0, 1, \dots, N-1$$

Vu que la fonction distance au centroïd $r(t)$ n'est invariante qu'à la rotation et la translation, les coefficients de Fourier doivent être normalisés afin qu'ils deviennent indépendants du changement d'échelle mais aussi du point de d'origine. À partir de la théorie de la transformée de Fourier, la forme générale des coefficients de Fourier de la fonction distance au centroïd $r(t)$, transformée à travers un changement d'échelle et un changement de point de départ à partir d'une fonction originale $r(t)^{(0)}$ est la suivante :

$$a_n = \exp(jn\tau) \cdot s \cdot a_n^{(0)}$$

Où a_n et $a_n^{(0)}$ sont les coefficients de Fourier de la transformé et de la forme originale respectivement, τ est l'angle généré par le changement de point de départ ; s est un facteur d'échelle. Considérons l'équation suivante :

$$b_n = \frac{a_n}{a_1} = \frac{\exp(jn\tau) \cdot s \cdot a_n^{(0)}}{\exp(j\tau) \cdot s \cdot a_1^{(0)}} = \frac{a_n^{(0)}}{a_1^{(0)}} \exp[j(n-1)\tau] = b_n^{(0)} \exp[j(n-1)\tau]$$

b_n et $b_n^{(0)}$ sont les coefficients normalisés de la forme transformée et de la forme originale respectivement. Si l'on ignore l'information phase et que l'on n'utilise que la magnitude des coefficients, alors $|b_n|$ et $|b_n^{(0)}|$ seront identiques. En d'autres termes $|b_n|$ sera invariant à la translation, la rotation, le changement d'échelle et le changement de point d'origine.

L'ensemble des magnitudes des coefficients de Fourier normalisés $\{|b_n|, 0 \leq n < N\}$ sera finalement utilisé comme descripteur dénoté de la sorte :

$$\{FD_n, 0 < n < N\}.$$

Le FD unidimensionnelle possède de « belles » propriétés telles qu'une dérivation et une normalisation assez simple mais aussi facile d'application. Comme indiqué dans [22] pour une reconnaissance efficace 10 FDs seront suffisants pour décrire la forme.

1.2.1.2 Méthodes structurelles

Avec l'approche structurelle, les formes sont découpées en segments de contour appelés *primitives*. Les méthodes structurelles diffèrent entre elles dans la sélection des primitives et dans la manière d'organiser celles-ci pour représenter la forme. Les méthodes de décomposition du contour les plus connues sont basées sur l'approximation polygonale, la décomposition de courbure ou encore sur l'ajustement des courbes « curve fitting »[23]. Le résultat de ceci sera encodé en une chaîne de la forme :

$$S = s_1, s_2, \dots, s_n.$$

Où le s_i peut être un élément d'une « *chain code* », un côté d'un polygone, un arc quadratique...etc. s_i peut contenir un certain nombre d'attributs comme la longueur, la courbure moyenne, la courbure maximale, l'énergie de flexion, l'orientation,...etc. La chaîne pourra par la suite être utilisée soit directement pour décrire la forme ou alors être une entrée pour un autre système de plus haut niveau d'analyse de forme. Dans ce qui suit nous allons présenter quelques exemples de cette famille de méthodes.

La « Chain Code » : la *chain code* décrit l'objet par une séquence de segments de ligne unitaires munis d'une orientation. Cette méthode a été introduite en 1961 par Freeman[24] qui présentait une méthode permettant d'encoder une configuration géométrique quelconque. Dans cette approche une courbe arbitraire est représentée par une séquence de petits vecteurs unitaires et un ensemble fini de directions possibles. Dans l'implémentation, une grille est superposée avec le contour de l'objet dans l'image binaire, chaque point du contour sera ensuite approximé avec le plus proche point de la grille, on obtiendra ainsi une image échantillonnée. D'un point de départ choisi préalablement, une *chain code* pourra être générée en concaténant une chaîne à 4 ou à 8 directions,

Figure 1.12. Une *chain code* à N directions ($N > 8$ et $N = 2^k$) est aussi possible, elle sera appelée une *chain code généralisée*.

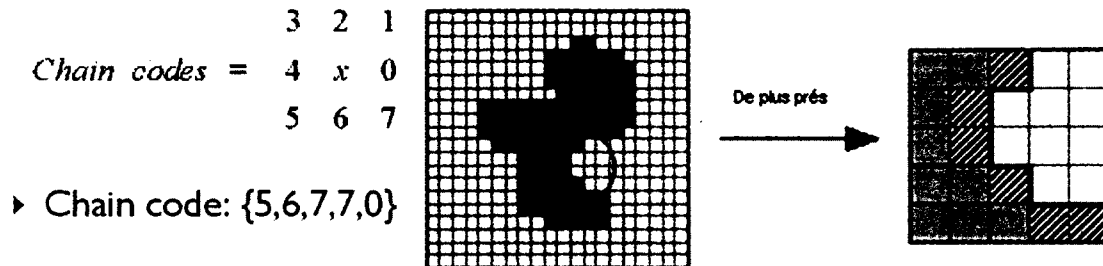


Figure 1.12 – Illustration de la représentation par *chain code*

Si la *chain code* est utilisée pour faire une mise en correspondance « *matching* » elle devra être indépendante du choix du point de départ sur le contour. Une possibilité, justement, pour la normaliser serait de trouver un pixel sur le contour qui minimiserait le nombre généré par la description -si celle-ci est exprimée en chiffres- ce point sera du coup considéré comme pixel de départ. Une autre alternative serait de représenter le contour par les différences entre les directions successives de la *chain code* au lieu des directions relatives. Ceci peut être obtenu en soustrayant chaque élément de la chaîne de son prédécesseur et en prenant le résultat modulo n , où n décrit la connexité, Figure 1.13. A la fin de ces opérations, une *chain code* invariante à la rotation peut être obtenue en effectuant une permutation cyclique, ce qui produit un nombre minimal. Mais une chaîne dérivée de cette manière n'est pas invariante au changement d'échelle. En plus du fait qu'une *chain code* possède généralement une très grande dimension et est sensible au bruit, elle est du coup utilisée comme entrée à un system plus élaboré d'analyse de forme comme l'approximation polygonale ou pour trouver la courbure du contour qui est un attribut perceptuel très important.

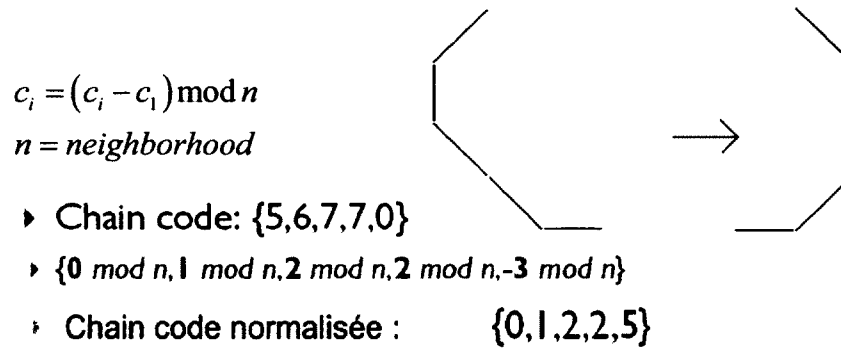


Figure 1.13 – Illustration de la *chain code* différentielle (normalisée)

- i) **Décomposition polygonale** : dans [25], le contour est découpé en segments de ligne par approximation polygonale. Ce seront donc les arêtes du polygone qui seront utilisées comme primitives (Figure 1.14). L'attribut pour chaque primitive sera ainsi exprimé sous forme d'une chaîne, en fonction de quatre éléments qui sont: l'angle interne, la distance au sommet suivant et ses coordonnées en x et en y . Clairement cet attribut ne sera pas invariant ni à la translation, ni à la rotation, ni au changement d'échelle. La similarité entre deux formes sera la distance d'édition [15] entre leurs deux chaînes d'attributs, ainsi et pour des raisons d'efficacité et de robustesse, seulement un nombre fixé d'attributs les plus pertinents sera sélectionné pour chaque forme. De la, une collection d'attributs appartenant à tous les modèles dans la base de données sera générée pour être utilisée comme index d'attributs. Ces attributs seront organisés en arbres binaires ou m -aires. Le *matching* entre formes deviendra dès lors un processus à deux étapes qui sont un *matching* attribut-à-attribut comme première étape ensuite un autre *matching* model-à-model.

Dans la première étape, et étant donné un attribut requête, celui-ci sera recherché dans l'arbre index, si un attribut particulier dans la base de données est trouvé comme étant similaire à la requête, la liste de modèles associée à cet attribut sera retrouvée. Dans la deuxième étape, le *matching* entre la requête et un model dans la liste sera basé sur la distance d'édition entre les deux chaînes de primitives.

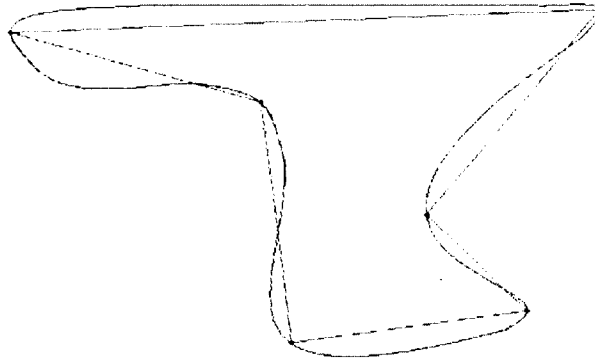


Figure 1.14 – Exemple de décomposition polygonale

Dans [26] Mehrotra et Gary représentent une forme comme une chaîne de vecteurs. Pour une forme donnée, un ensemble de points d'intérêt sont détectés à partir de l'approximation polygonale du contour. Etant donnée ces n points d'intérêt, une paire en est choisie pour former un vecteur de base. Celui-ci est normalisé en un vecteur unitaire le long de l'axe des x . Tous les autres points d'intérêt de la forme seront transformés à ce nouveau système de coordonnées. La forme est alors représentée par les coordonnées $(x_1, y_1), \dots, (x_n, y_n)$ ou les (x_i, y_i) sont les coordonnées normalisées du $i^{\text{ème}}$ point d'intérêt. (Figure 1.15)

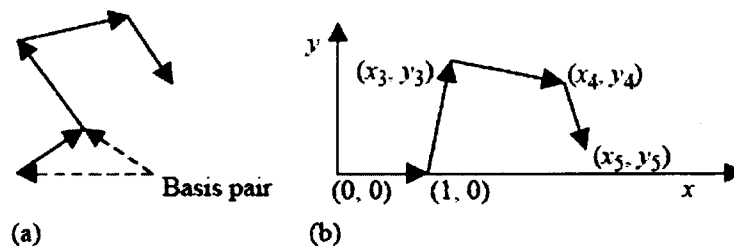


Figure 1.15 – Vecteurs chaîne: (a) description originale de la frontière; (b) Description normalisée

Quatre paramètres seront rajoutés à l'ensemble de coordonnées pour créer le vecteur d'attributs final, ces quatre sont : S , T_x , T_y , θ , qui représentent respectivement l'échelle, la translation et l'angle du vecteur de base. La similarité entre attributs sera

mesurée par la distance euclidienne. Les vecteurs d'attributs de contour seront organisés dans un B -arbre de dimension k . Pendant une requête et pour un attribut donné, une recherche est effectuée dans l'index et une liste de formes ayant des attributs similaires est produite. Le *matching* d'un ou plusieurs attributs ne garantit pas, par contre, une similitude complète de la forme, ce qui fait qu'une fois que des formes avec des attributs similaires sont retrouvées, la similarité entre formes est vérifiée en juxtaposant chaque forme trouvée avec la requête et en évaluant le degré d'imbrication entre elles.

- ii) **Décomposition selon la courbure :** Berretti et al. [27] ont étendu le modèle utilisé dans [25] pour une reconnaissance de formes plus générales. Dans [27], les points de *zero-crossing* de la courbure du contour lissé par une gaussienne, ces points disons nous, seront utilisés pour obtenir les primitives appelées dans ce cas « *jetons* » (Figure 1.16). L'attribut pour chaque jeton sera sa courbure maximale et son orientation, la similarité entre deux jetons, dans ce cas, sera mesurée par une distance euclidienne pondérée. Le fait que l'attribut inclut une orientation, il ne sera plus invariant à la rotation.

Un arbre M -aire sera exploité pour indexer les jetons dans la base de données des attributs (Figure 1.17). Etant donné une forme requête, le processus de recouvrement de formes similaires de la base de données prend deux étapes : la première de ces étapes est de retrouver des jetons similaires. Pour tous les N jetons de la forme requête, des jetons similaires seront sortis en parcourant l'index (N fois). L'ensemble des jetons en sortie ayant le même identifiant de forme, constituera une forme similaire potentielle. La deuxième étape consistera à faire correspondre la forme requête avec cette forme similaire potentielle, en utilisant un algorithme de *matching* model-à-model qui prendra $O(MN)$ opérations (M et N sont, respectivement, le nombre de jetons des deux formes à faire correspondre). Le *matching* des jetons dans les deux étapes implique un seuillage qui peut se faire de manière ad hoc ou par des méthodes empiriques.

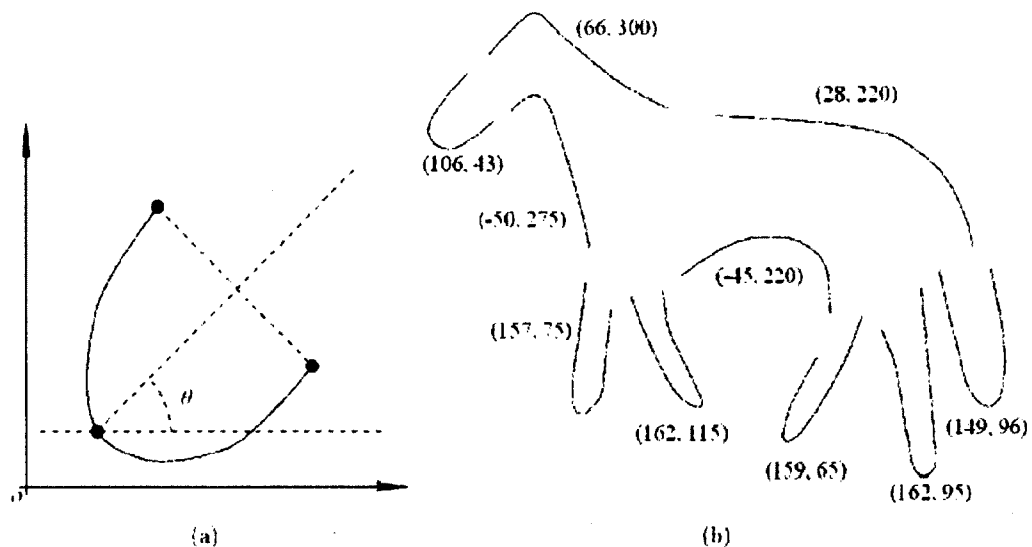
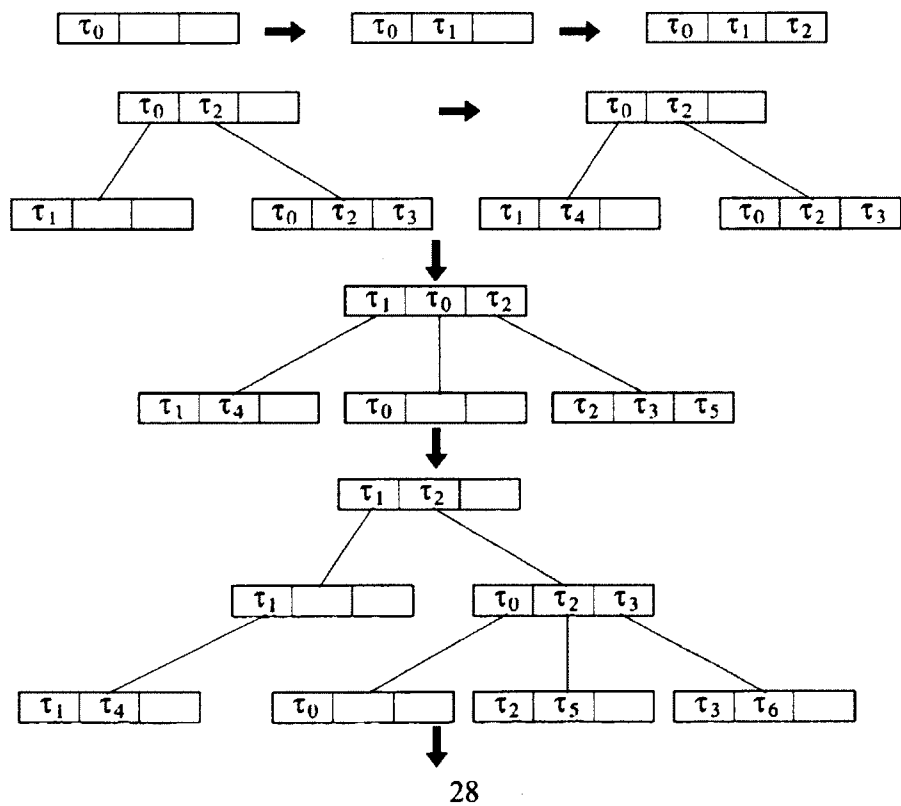


Figure 1.16 – (a) est l'orientation du jeton ; (b) Une forme de cheval décomposée en plusieurs jetons. Les nombres correspondants à chaque jeton sont la courbure et l'orientation de celui-ci



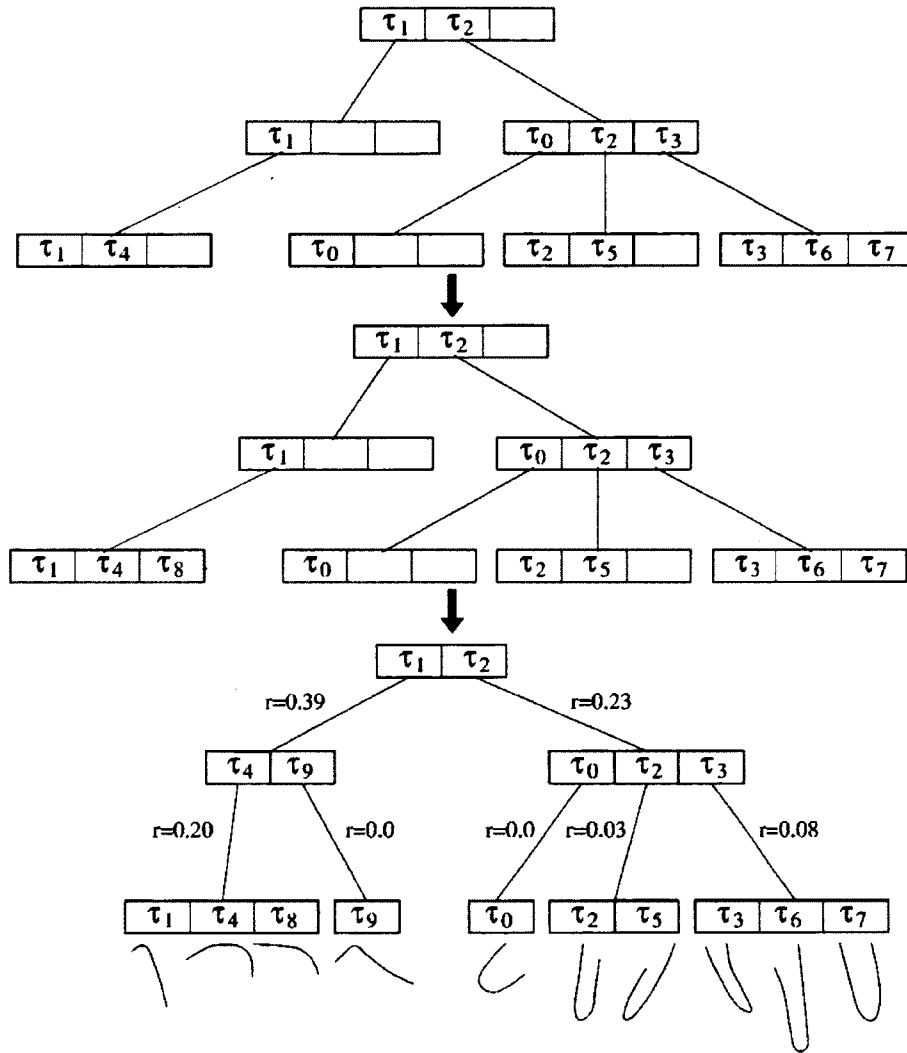


Figure 1.17 – Indexation des jetons dans un arbre *M-aire*

Vu que l'arbre est parcouru un nombre significatif de fois, il n'est pas sur que l'indexation des jetons est meilleure qu'un *matching* model-à-model direct. L'efficacité du *matching* est très dépendante du nombre de jetons pour chaque forme, mais aussi du degré du lissage effectué en amont.

1.2.2 Méthodes basées sur la région

Dans les techniques basées sur la région, tous les points dans la forme seront pris en compte dans l'obtention de la représentation de celle-ci, au lieu de juste ceux du contour comme dans les techniques basées sur le contour. On peut citer comme exemples de méthodes basées sur la région, les moments de région, la matrice de forme, le *Convex Hull*, l'axe médian,...Etc. Tout comme les méthodes basées sur le contour, celles basées sur la région peuvent aussi être divisées en globales ou structurelles, selon qu'elles séparent la forme en sous-parties ou pas.

1.2.2.1 Méthodes globales

Ces méthodes traitent la forme comme un tout, la représentation résultante est un vecteur d'attributs numériques qui sera utilisé pour décrire la forme. La similarité entre formes est mesurée par des distances métriques entre vecteurs.

- i) **Les moments région :** Parmi les descripteurs de forme basés sur la région, les moments sont très populaires. Ils incluent les moments invariants, les moments de Zernike, les moments algébriques, les moments radiaux de Chebyshev, etc. La forme générale de la fonction moment m_{pq} d'ordre $(p+q)$ d'une région, peut être donnée par :

$$m_{pq} = \sum_x \sum_y \Psi_{pq}(x, y) f(x, y) \quad p, q = 0, 1, 2, \dots$$

où Ψ_{pq} est nommé *noyau de pondération du moment* ou encore *ensemble de base* ; $f(x, y)$ est l'équation de la région (eq1.1).

- **Les moments invariants (IM)** aussi appelés moments géométriques ont pour ensemble base $\Psi_{pq} = x^p y^q$ et leur fonction m_{pq} d'ordre $(p+q)$ est donnée par :

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad p, q = 0, 1, 2, \dots$$

Les moments géométriques invariants à la translation seront définis comme :

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad p, q = 0, 1, 2, \dots$$

Ou $\bar{x} = m_{10}/m_{00}$ et $\bar{y} = m_{01}/m_{00}$, un ensemble de 7 moments invariants sera donné par :

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &\quad \cdot [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} + \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}^2(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &\quad \cdot [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

où $\eta_{pq} = \mu_{pq} / \mu_{00}^\gamma$ et $\gamma = 1 + (p+q)/2$ pour $p+q = 2, 3, \dots$

Les IM sont très simples et rapides à calculer, en plus du fait qu'ils soient invariants à la rotation, la translation et au changement d'échelle. Ils présentent, par contre, certains inconvénients dus à leur construction même [28] :

- Les moments d'ordre supérieur sont très sensibles au bruit.
- Vu que l'ensemble base des IM n'est pas orthogonal, ceux-ci souffrent d'un niveau élevé de redondance d'informations.
- Une grande variation dans l'intervalle des valeurs: vu que la base contient une fonction de puissance (de p et q), les moments calculés présenteront une grande variation dans l'intervalle dynamique des valeurs pour des ordres différents.

- **Les moments algébriques** : ils sont calculés à partir des m premiers moments centraux, et sont donnés sous forme de valeurs propres de matrices prédéfinies $M_{[j,k]}$. Les valeurs de cette matrice sont des facteurs des moments centraux. Les moments algébriques peuvent être construits pour des ordres aléatoires et sont invariants aux transformations affines. Ils sont, par contre, très dépendants du type de forme (ils peuvent donner de très bons résultats pour certaines configurations de formes et des résultats beaucoup plus médiocres pour d'autres).
- **Les moments de Zernike (ZM)** : ce sont des moments orthogonaux, dérivés des polynômes orthogonaux du même nom [14] :

$$V_{nm}(x, y) = V_{nm}(r \cos \theta, r \sin \theta) = R_{nm}(r) \exp(jm\theta)$$

où $R_{nm}(r)$ est le polynôme orthogonal :

$$R_{nm}(r) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \times \left(\frac{n-2s+|m|}{2} \right)! \left(\frac{n-2s-|m|}{2} \right)!} r^{n-2s}$$

$n = 0, 1, 2, \dots; 0 \leq |m| \leq n$; et $n - |m|$ est pair.

Les polynômes de Zernike sont tout un ensemble de fonctions à valeurs complexes, orthogonales sur le disque unitaire $x^2 + y^2 \leq 1$. Le moment de Zernike d'ordre n et de répétition m de la région $f(x, y)$ d'une forme sera donné par :

$$Z_{nm} = \frac{n+1}{\pi} \sum_r \sum_\theta f(r \cos \theta, r \sin \theta) r \cdot R_{nm}(r) \cdot \exp(jm\theta) \quad r \leq 1$$

Ces moments de Zernike présentent certains avantages dont:

- Les magnitudes des moments de Zernike sont invariantes à la rotation.

- Ils sont assez robustes au bruit mais aussi aux variations mineures dans la forme.
- Ils présentent une redondance d'informations minimale car la base est orthogonale.

Les moments de Zernike (moments orthogonaux et continus) ont cependant quelques lacunes, on peut citer :

- L'espace des coordonnées de l'image doit être transformé en un domaine où le polynôme orthogonal est défini (disque unitaire).
- La complexité de calcul du polynôme de Zernike augmente très rapidement quand on passe à des ordres supérieurs.
- Les intégrales continues devront être approximées par des sommations ce qui, non seulement produit des erreurs numériques dans les moments calculés, mais aussi affecte les propriétés analytiques telles que l'invariance à la rotation et l'orthogonalité.

ii) Descripteur de Fourier générique : Bien qu'ayant une performance assez robuste les moments de Zernike présentent certains inconvénients. En plus de ceux déjà cités, on remarque aussi que les attributs « capturés » par les moments de Zernike ne sont pas consistants, dans le sens où les attributs radiaux sont dans le domaine spatial et les attributs circulaires sont, eux, dans le domaine spectral, ce qui exclut l'analyse multi-résolution d'une forme dans la direction radiale. Ensuite, les attributs circulaires spectraux ne sont pas capturés équitablement à chaque ordre. Ce qui peut causer une perte d'attributs pertinents dans certains ordres.

Pour contourner ces inconvénients, Zhang et Lu [29] ont proposé un descripteur de Fourier générique (GFD). Celui-ci est acquis en appliquant une transformée de Fourier à 2-D sur une image à coordonnées polaires échantillonnée Figure 1.18:

$$PF_2(\rho, \phi) = \sum_r \sum_i f(r, \theta_i) \exp \left[j2\pi \left(\frac{r}{R} \rho + \frac{2\pi i}{T} \phi \right) \right]$$

Ou $0 \leq r \leq R$ et $\theta_i = i(2\pi/T)$ ($0 \leq i < T$); $0 \leq \rho < R$, $0 \leq \phi < T$. R et T sont, respectivement, la résolution à fréquence radiale, et celle à fréquence angulaire. Enfin, les coefficients normalisés formeront le GFD. La similarité entre deux formes sera calculée par la distance *City Block* entre leurs GFD respectifs.

En comparant le GFD aux moments de Zernike, il s'avère que : le GFD est plus simple à calculer, les attributs sont purement spectraux et montrent de meilleurs performances de recouvrement car l'analyse multi-résolution dans les deux directions radiale et spectrale, est possible. Zhang et Lu ont aussi montré que le GFD donnait de meilleurs performances que le DF, les moments géométriques et les moments de Zernike.

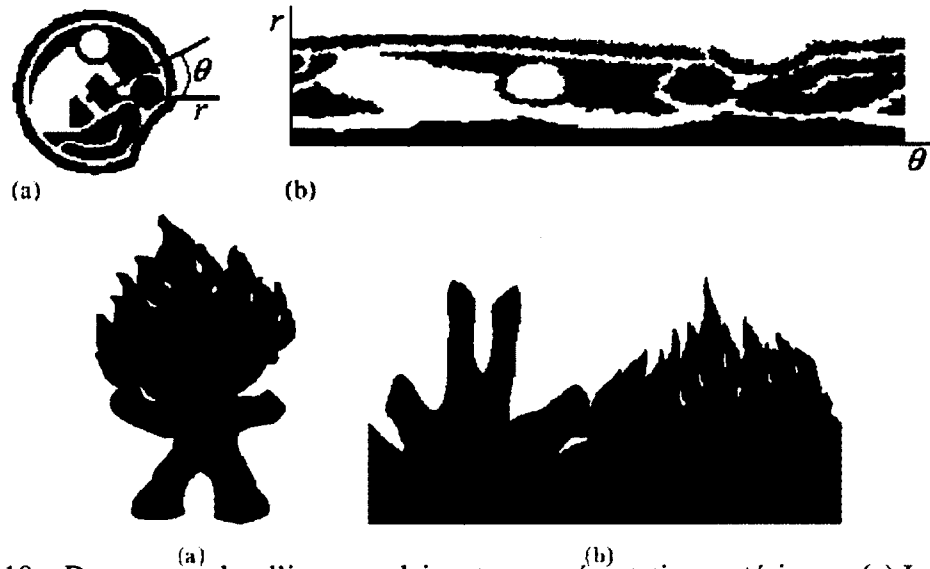


Figure 1.18 – Deux exemples d'image polaire et sa représentation cartésienne. (a) Image originale à coordonnées polaires ; (b) L'image polaire dans (a) représentée dans un espace cartésien.

iii) **Matrices de formes** : Les méthodes de description de forme qui utilisent un échantillonnage de l'image pour acquérir les informations de la forme, usent en général de « grilles » rectangulaires à cette fin d'échantillonnage. Mais les représentations dérivées de cette manière ne sont souvent pas invariantes ni à la translation, ni à la rotation, ni au changement d'échelle, ce qui engendrera des normalisations supplémentaires pour y palier. Goshtasby propose l'utilisation d'une *matrice de forme* dérivée d'un échantillonnage à l'aide d'une grille circulaire [30]. L'idée est similaire à un échantillonnage classique, cependant, au lieu de superposer une grille carrée(ou rectangulaire), une grille à coordonnées polaires formée de cercles concentriques et de lignes (diamètres), sera superposée sur le centre de masse de la forme (Figure 1.19(a)). Les valeurs binaires de la forme seront prélevées aux intersections des diamètres et des cercles. La matrice sera remplie de sorte que les cercles correspondent aux colonnes et les diamètres aux lignes. On note que l'échelle de la forme sera normalisée en utilisant son diamètre maximal, et ce en amont de l'échantillonnage. La matrice résultante sera invariante à la translation, à la rotation ainsi qu'au changement d'échelle. Vu que la densité d'échantillonnage n'est pas constante avec une grille polaire, Taza et Suen représentent la forme avec une matrice pondérée, voir [31] pour plus de détails.

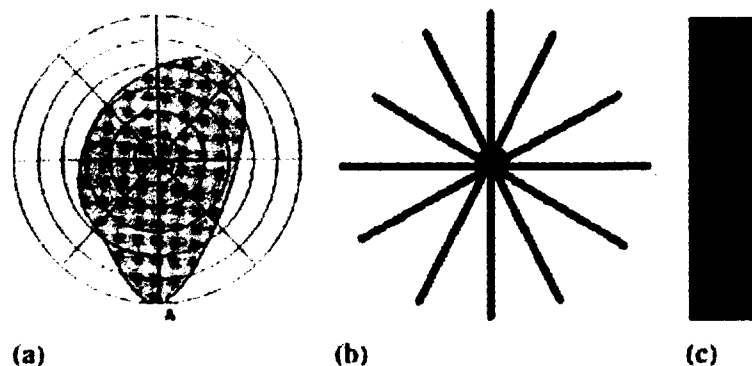


Figure 1.19 – (a) Echantillonnage en coordonnées polaires ; (b) Une étoile formée de barres
 (c) Forme rectangulaire

Le fait que la matrice ne soit qu'un échantillonnage épars de la forme, elle sera facilement affectée par le bruit (comme toutes les méthodes de cette famille). Parui et al. proposent une description de forme basée sur les surfaces relatives de la forme contenues dans les anneaux concentriques au centre de masse [2]. Soient : L le diamètre maximal d'une forme S à décrire, C_k le k^{ieme} anneau des n anneaux concentriques, obtenus en sectionnant le diamètre maximale L en n segments égaux. Un ratio de surface invariant sera défini de la sorte :

$$x_i = \frac{A(S \cap C_i)}{A(C_i)}$$

Ou $A(\cdot)$ est la fonction surface. Le descripteur de forme sera, du coup, le vecteur d'attributs $x = [x_1, \dots, x_n]^T$. Mais bien qu'étant plus compacte et plus robuste que la matrice de forme, le ratio de surface ignore les altérations du motif à l'intérieur des anneaux mesurés ; Conséquemment, les deux formes dans Figure 1.19(b) et Figure 1.19(c), seront les mêmes selon ce descripteur. Ce problème peut être résolu en utilisant la transformée spectrale discutée dans le GFD.

- iv) **Descripteurs topologiques** : Cette catégorie de descripteurs s'intéresse aux attributs d'une forme qui sont invariants à des transformations dites *rubber-sheet*, c'est à dire des mouvements d'étirements ou de rétrécissement de cette forme du moment que ceux-ci ne causent pas de coupure de la forme ou de fusion de deux composantes séparées au départ. Un descripteur topologique très connu est le « *Nombre d'Euler* » E , donné par la différence entre le nombre de composantes connexes C et le nombre de trous H dans la forme : $E = C - H$. Dans Figure 1.20 par exemple,



Figure 1.20 – Trois formes avec des nombres d'Euler différents.

le nombre d'Euler pour ces trois formes sera, respectivement: 1, -1 et 0. L'homologie est, à ce propos, très adaptée comme outil de calcul des propriétés topologiques comme à travers les nombres de *Betti* ou groupes d'homologie – décrits avec plus de détails plus loin dans le mémoire- qui donnent, outre le nombre de composantes connexes et le nombre de trous dans l'objets, le nombre de cycles (boucles) dans l'objet. En étudiant l'évolution de ceux-ci au fil de l'apparition de points critiques on peut déduire des propriétés topologique de l'objet.

1.2.2.2 Méthodes structurelles

Tout comme les méthodes structurelles basées sur le contour, celles basées sur la région décomposent la forme en parties qui seront utilisées par après pour la description et la représentation de la forme.

- i) **Enveloppe convexe** : Une région R est dite convexe si et seulement si, pour tout deux points $x_1, x_2 \in R$, l'intégralité du segment x_1x_2 est à l'intérieur de la région. L'*enveloppe convexe* d'une région R est la plus petite région convexe H contenant celle-ci et satisfaisant la condition $R \subset H$, la différence $H - R$ est appelée *déficiences convexe* D de la région R . Il est commun d'effectuer un lissage du contour de la forme avant le partitionnement et ce pour diminuer l'effet du bruit.

La représentation de la forme peut être obtenue par un processus récursif, ce qui résulte en un arbre de concavité Figure 1.21. L'enveloppe convexe d'un objet est d'abord trouvée avec ses déficiences convexes, ensuite les enveloppes convexes et les déficiences, des déficiences elles-mêmes sont trouvés et ainsi de suite, jusqu'à ce que toutes les déficiences soient convexes.

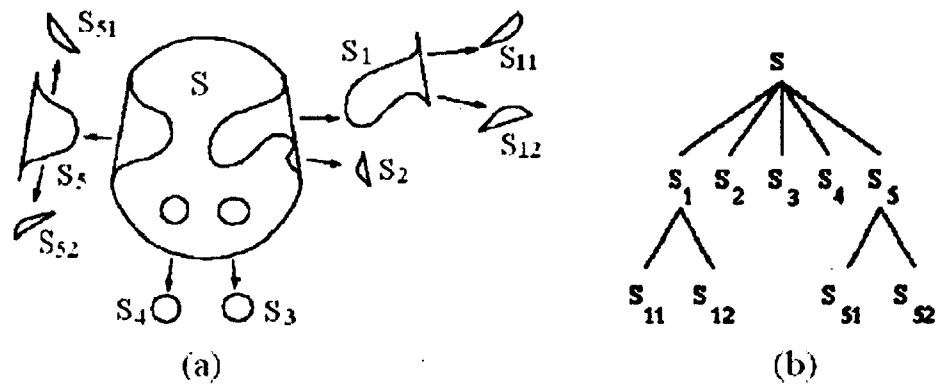


Figure 1.21 – (a) Enveloppe convexe et ses concavités ; (b) Représentation en arbre des concavités de l'enveloppe.

Chaque concavité est décrite par sa superficie, sa longueur de corde (la ligne reliant les deux bouts de la coupure de la concavité), sa courbure maximale et la distance entre le point de courbure maximale et la corde. La mise en correspondance entre deux formes devient une mise en correspondance de chaînes ou de graphes. La représentation à l'aide d'une enveloppe convexe est invariante à la translation, la rotation et au changement d'échelle. Elle est robuste aux contours bruités (après le filtrage), c'est, cependant, l'extraction de l'enveloppe qui reste problématique car très coûteuse en temps de calcul.

- ii) **L'axe médian** : Tout comme l'enveloppe convexe, le squelette d'une région peut aussi être employé pour représenter et décrire une forme. Un squelette peut être défini comme étant un ensemble d'axes médians connectés, le long des branches d'une figure [18]. Dans le cas, par exemple d'un caractère manuscrit épais, le squelette sera le chemin prit par le stylo. En fait l'idée de base du squelette, est d'éliminer les informations redondantes et de ne garder que l'information topologique concernant la structure de l'objet qui nous est utile lors de la reconnaissance. La méthode du squelette est représentée par la transformée de l'axe médian (*Medial Axis Transform*) de Blum [9]. L'axe médian est le locus des centres des disques compris entièrement dans la forme. Ceci est illustré dans la Figure 1.22.

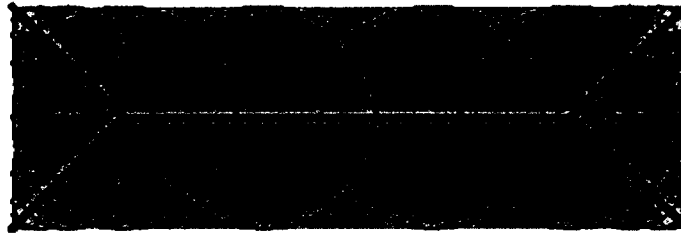


Figure 1.22 – Construction de l'axe médian d'une forme rectangulaire

La ligne épaisse dans la figure constitue le squelette de la forme rectangulaire grisée. Le squelette pourra ensuite être décomposé en segments et représenté par un graphe selon certains critères. La mise en correspondance de formes devient une mise en correspondance entre graphes. Le calcul de l'axe médian est très complexe, en plus du fait qu'il est très sensible au bruit sur le contour de la forme.

1.3 Mesures de similarité

Les mesures de similarité sont un ingrédient essentiel à considérer lorsqu'on souhaite traiter de la description, l'analyse et la reconnaissance de forme. On observe deux directions de recherche dans le domaine des mesures de similarité, d'une importance égale. La première explore l'efficacité de recouvrement, en termes de résultats concordants ou pas avec les attentes de l'utilisateur (mesurée par le Précision/rappel). Ainsi, on voudra notre mesure la plus riche et globale possible et limitée le moins possible par des propriétés restrictives. La deuxième direction, quant à elle, veut que la mesure soit la plus concise (rapide) possible, car le nombre d'objets dans la base de données peut atteindre des proportions très considérables et en même temps les mesures de similarité sont souvent très coûteuses en temps de calcul [14,32,33]. Il y a clairement donc un compromis à faire selon ces deux termes.

Une mesure de similarité doit concorder le plus possible avec la perception visuelle, ce qui implique les propriétés idéales suivantes :

- Elle doit être universelle dans les sens où elle doit distinguer des objets de forme arbitraires n'ayant aucune restriction à priori.
- Elle doit permettre la reconnaissance d'objets perceptuellement similaires même s'ils sont mathématiquement non identiques.
- Elle doit passer outre les distorsions telles que : bruit de numérisation, erreurs de segmentation où occultations partielles.
- Elle doit être indépendante de l'échelle, l'orientation où la position de l'objet, mais aussi des transformations affines et distorsions de perspective.

En mathématiques, les mesures de similarité et de dissimilarité sont définies comme suit: Soit un objet model $O \in U$ ou U est un model universel. Soit $s: U \times U \mapsto R$ une mesure de similarité, ou $s(O_i, O_j)$ est considéré comme le score de la similarité des deux objets O_i et O_j . Mais dans la majorité des cas il est préférable d'utiliser une mesure de dissimilarité $d: U \times U \mapsto R$ équivalente à s tel que : $s(Q, O_i) > s(Q, O_j) \Leftrightarrow d(Q, O_i) < d(Q, O_j)$.

Une mesure de similarité assigne la plus petite distance (ou le plus grand score) aux objets les plus similaires et vice-versa. En général, une mesure métrique est nécessaire pour organiser efficacement les éléments d'une base de données à travers des vecteurs multidimensionnels. L'espace de distances métriques est défini suivant [27] comme: un espace \mathbb{R}^N est un espace métrique si pour toute paire d'éléments x et y , il existe une distance $d(x, y)$ qui satisfait les propriétés suivantes :

- $d(x, y) \geq 0$ (positivité)
- $d(x, y) = 0$ si et seulement si $x=y$ (identité).
- d est continue.

- Une différence majeure est supposée causer une plus grande dissimilarité que quelques autres différences mineures.
- $d(x, z) \leq d(x, y) + d(y, z)$ (inégalité triangulaire).

En supposant que $A(a_0, a_1, \dots, a_{n-1})$ et $B(b_0, b_1, \dots, b_{n-1})$ soient respectivement le vecteur requête et le vecteur cible, il existe plusieurs méthodes classiques pour calculer le vecteur distance dans l'espace métrique.

Distance de Minkowski : cette distance est basée sur la norme L_p et est définie de la sorte :

$$d_p(A, B) = \left(\sum_{i=0}^{N-1} |a_i - b_i|^p \right)^{1/p}$$

Quand $p=1$, $d_p(A, B)$ est une L_1 distance de norme-1. Elle est aussi connue comme la distance « city block » ou encore « distance de Manhattan ». La distance de Manhattan d_{Man} entre deux vecteurs A et B , comme présenté dans Figure 1.23(a), est donnée par :

$$d_{Man}(A, B) = \sum_{i=0}^{N-1} |a_i - b_i|$$

Quand $p=2$, $d_2(A, B)$ sera alors la distance euclidienne d_{Eu} Figure 1.23(b).

$$d_{Eu}(A, B) = \sqrt{\sum_{i=0}^{N-1} (a_i - b_i)^2}$$

Quand $p \rightarrow \infty$ on obtient la distance à norme infinie L_∞ aussi appelée distance de *Chebyshev*. Cette distance est une métrique définie sur un espace vectoriel ou la distance entre deux vecteurs est le maximum de leurs distances selon n'importe qu'elle dimension de coordonnées. Figure 1.23(a).

$$L_\infty(A, B) = \max_{0 \leq i \leq N-1} \{|a_i - b_i|\}$$

Distance cosinus :

On calcule ici la différence dans la direction quelle que soit la longueur des vecteurs donnés. La différence sera donc donnée par l'angle entre les deux vecteurs.

$$d_{\cos}(A, B) = 1 - \cos \theta = 1 - \frac{A \cdot B^T}{\|A\| \cdot \|B\|} = 1 - \frac{\sum_{i=0}^{N-1} a_i b_i}{\sqrt{\sum_{i=0}^{N-1} a_i^2} \sqrt{\sum_{i=0}^{N-1} b_i^2}}$$

Etant donné que cette distance ne prend en compte que l'angle, deux vecteurs ayant la même direction auront la même distance cosinus de n'importe quel autre vecteur. Figure 1.23(c).

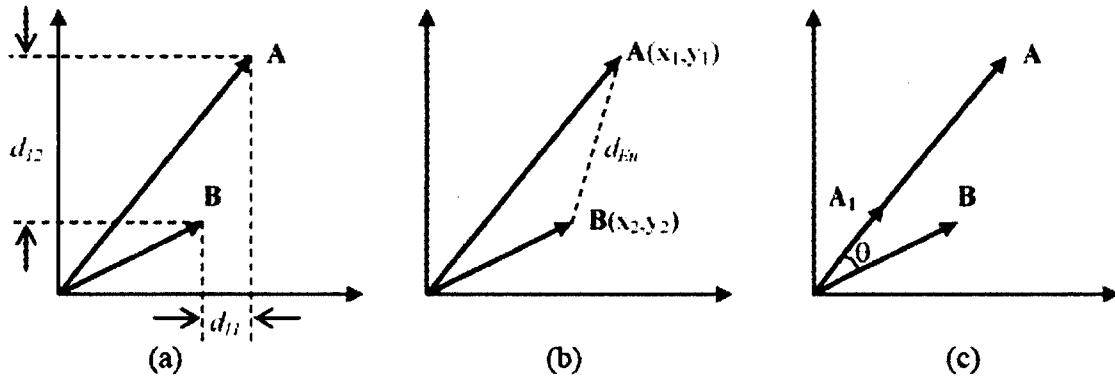


Figure 1.23 (a) La distance de Manhattan $d_{Man} = d_{11} + d_{12}$ et la distance à norme infinie $L_{\infty} = \max(d_{12}, d_{11}) = d_{12}$; (b) Distance Euclidienne $d_{Eu} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$; (c) les vecteurs A et A_1 ont la même distance cosinus par rapport au vecteur B .

Distance statistique χ^2 : Elle est définie de la sorte

$$D_{\chi^2}(A, B) = \frac{1}{2N} \sum_{i=0}^{N-1} \frac{(a_i - b_i)^2}{(a_i + b_i)}$$

Elle mesure le peu de probabilité qu'un élément d'un vecteur soit pris de la population décrite par l'autre vecteur.

Divergence de Kullback-leiber (KL) :

$$D_{KL}(A, B) = \sum_{i=0}^{N-1} a_i \log \frac{a_i}{b_i}$$

Sa principale qualité est qu'elle est assez robuste au bruit.

Divergence de Jeffrey (JD) :

$$D_{JD}(A, B) = \sum_{i=0}^{N-1} \left(a_i \log \frac{a_i}{b_i} + b_i \log \frac{b_i}{a_i} \right)$$

Elle est symétrique en comparaison avec *KL*.

Distance de la forme quadratique (QF) :

La distance de la forme quadratique est définie, selon la similarité des éléments des vecteurs, de la sorte :

$$D_{QF}(A, B) = \sqrt{(A - B) \cdot S \cdot (A - B)^T}$$

Ou $S = [s_{ij}]$ est la matrice de similarité, s_{ij} décrit la similarité (distance basique) entre les éléments a_i et b_i . Certaines de ces distances basiques sont définies comme ceci :

a) $s_{ij} = 1 - d_{ij} / d_{\max}$, ou d_{ij} est la distance entre a_i ($i^{\text{ème}}$ élément de A) et b_i ($i^{\text{ème}}$ élément de B), d_{\max} est le maximum des d_{ij} .

b) $s_{ij} = \exp\left(-\sigma \left(d_{ij} / d_{\max}\right)^2\right)$, ou σ est une constante positive, et d_{ij} reste le même que dans (a).

Distance de Mahalanobis :

La distance de *Mahalanobis* prend en compte la covariance des variables dans le calcul des distances. Avec cette mesure, les problèmes d'échelle et de corrélation inhérents à la distance *Euclidienne* sont éliminés. Pour comprendre comment ceci est effectué, considérons le fait

que lorsque la distance Euclidienne est utilisée, l'ensemble des points équidistants d'un point donné forme une sphère. La distance de *Mahalanobis* étire cette sphère pour corriger les échelles respectives des différentes variables et pour tenir compte de la corrélation entre ces variables.

La distance de *Mahalanobis* entre deux vecteurs A et B d'une même distribution et qui a une covariance décrite par la matrice C , cette distance, disons nous, est donnée par :

$$d_{Man} = \sqrt{(A - B) \cdot C^{-1} \cdot (A - B)^T}$$

Distance de Hausdorff :

En topologie, la distance de Hausdorff mesure l'éloignement de deux sous-ensembles d'un espace métrique. La distance de Hausdorff entre les deux vecteurs A et B est définie comme étant le plus petit nombre réel r tel que le r -voisinage de A contienne B et le r -voisinage de B contienne A . Dans la pratique, la distance de *Hausdorff* entre A et B est la plus grande distance existante entre un point aléatoirement prit dans A et le point correspondant de B le plus proche, i.e. en choisissant un point dans A , la distance entre ce point et le point le plus proche dans B est inférieure ou égale à la distance de *Hausdorff* entre A et B . Figure 1.24.

La fonction de cette distance est donnée par :

$$d_{Hau}(A, B) = \max \left\{ \max_{b \in B} \left[\min_{a \in A} (d(a, b)) \right], \max_{a \in A} \left[\min_{b \in B} (d(a, b)) \right] \right\}$$

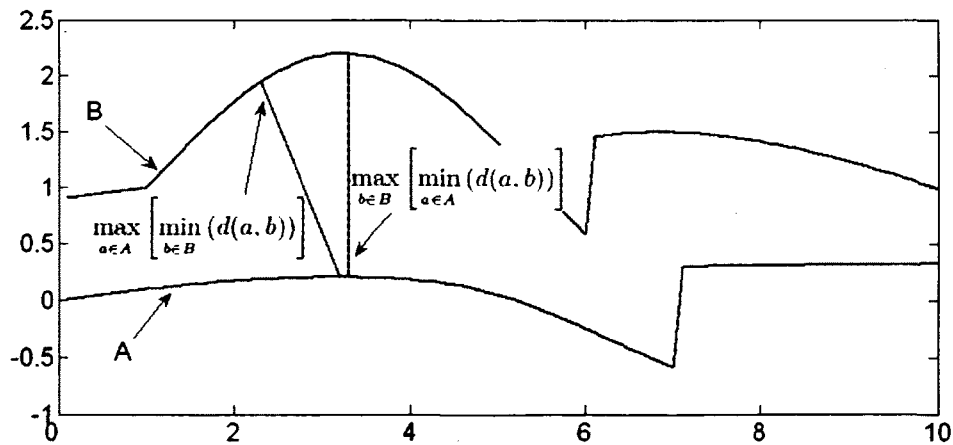


Figure 1.24 – Distance de Hausdorff

La distance de *Hausdorff* n'est pas invariante aux transformations affines, elle est aussi assez sensible au bruit. Néanmoins, l'avantage de son utilisation réside dans le fait qu'il est possible de l'utiliser dans des mises en correspondance (*matching*) partielles.

Distance d'édition:

La distance d'édition entre deux objets est définie comme étant le nombre d'opérations de base (insertion, suppression et substitution) nécessaires pour transformer une représentation en une autre [14,34]. Elle est calculable entre deux ensembles ordonnés de vecteurs. La différence entre formes sera ainsi calculée en termes de coûts d'édition. Pour mesurer la distance d'édition, on utilise les méthodes de distance de *Hamming*, de *Levenshtein*, de *Damerau-Levenshtein* et de *Jaro-Winkler*.

1.4 Evaluation des résultats

Pour évaluer la performance des différents algorithmes de reconnaissance de forme, une mesure de performance efficace est nécessaire. Ce type de mesures est généralement basé sur des statistiques extraites de tests subjectifs. Des mesures de performance différentes utilisent souvent des tests subjectifs différents, ce qui donne des définitions de précision de recouvrement différentes. Nous présentons dans cette section les mesures d'efficacité de recouvrement et de reconnaissance communément utilisées.

1.4.1 La paire *Précision/Rappel*

Le *Précision/Rappel* est probablement la mesure de performance la plus répandue dans la littérature. Pour chaque requête le système renvoi une liste pondérée. Chaque élément de cette liste est marqué comme étant pertinent ou pas selon le résultat du test. L'efficacité sera par la suite calculée par le *Précision/Rappel*. Le *Rappel* ou *rappel* mesure la capacité à retrouver des éléments pertinents à partir de la base de données. Il est défini par le ratio du nombre d'éléments pertinents retrouvés sur le nombre d'éléments pertinents dans la base de données. La précision est définie par le ratio du nombre d'éléments pertinents retrouvés au nombre total d'éléments retourné.

$rappel = \#(\text{éléments pertinents retournés}) / \#(\text{éléments pertinents dans la base de données})$

$précision = \#(\text{éléments pertinents retrouvés}) / \#(\text{éléments retrouvés})$

La fonction $\#(\alpha)$ donne le nombre d'éléments de α .

En notant le nombre d'éléments pertinents retrouvés par a , le nombre d'éléments non pertinents retrouvés par b et le nombre total d'éléments pertinents dans la base, non-retrouvés par c , on aura :

$$rappel = \frac{a}{a+c} \quad \text{et} \quad précision = \frac{a}{a+b}$$

Il sera évident que $rappel \in [0,1]$ et $précision \in [0,1]$

La précision diminue lorsque le rappel augmente et ce car dans le processus de vouloir retrouver tous les éléments pertinents pour une requête, quelques éléments non pertinents se glissent parmi les résultats. On utilise ainsi le rappel et la précision de paire pour indiquer l'efficacité du système.

1.4.2 Pourcentage de résultats pondérés (PWH)

Le PWH est comparable au rappel présenté dans 1.4.1 à la seule différence qu'ici au lieu de noter un élément comme étant pertinent ou pas (ce qui s'apparente à une approche binaire), on lui assigne plutôt un poids de pertinence w_i et ce pour chaque élément de l'ensemble de données [35]. Le poids w_i est obtenu manuellement de la façon suivante : étant donné un groupe d'utilisateurs et une base de données et en assignant initialement un poids nul à tous les éléments de la base, à chaque fois qu'un utilisateur choisit un certain ensemble d'éléments qui ressemblent à une certaine requête, les poids de ces éléments seront incrémentés de un. De cette manière, le poids w_i d'un élément est égal au nombre d'utilisateurs l'ayant choisi comme ressemblant à la requête.

Pour chaque requête, un nombre fixe de résultats est sorti par le système. Le pourcentage de bons résultats pondérés P sera défini de la sorte :

$$P = \frac{\sum_{i=1}^n w_{\pi(i)}}{\sum_{j=1}^N w_j}$$

où n est le nombre d'éléments retrouvés par le système ; $\pi(i)$ décrit les éléments retrouvés par un test ; N est le nombre total d'éléments dans la base de données ; $w_{\pi(i)}$ et w_j sont les poids donnés par les utilisateurs aux éléments qu'ils considèrent comme pertinents pour la requête. Pour produire la moyenne de performance du système, plusieurs requête seront lancées et la moyenne des pourcentages de bons résultats pondérés est calculée et sera notre indicateur de l'efficienne du système.

1.4.3 Pourcentage du classement selon la similarité (PSR)

Dans le *percentage of similarity ranking* (PSR), chaque utilisateur va assigner un rang ou une note à chaque élément dans l'ensemble de données selon le degré de sa ressemblance à la requête, au lieu de le noter simplement comme pertinent/non-pertinent comme dans le cas du RPP, où alors en n'incrémentant son poids que s'il est choisi par un des utilisateurs comme dans le PWH. Le résultat final de ce processus de test initial sera une matrice $\{Q_i(i, k)\}$, où $Q_i(i, k)$ indique le pourcentage d'utilisateurs qui ont rangé le $i^{ème}$ élément à la $k^{ième}$ position par rapport à la requête j . À partir de $Q_i(i, k)$, on pourra calculer la valeur moyenne $\bar{p}_j(i)$ et l'écart type $\sigma_j(i)$, ces deux mesures représentant le rang moyen du $i^{ème}$ élément pour la requête j et une mesure du consensus des utilisateurs sur le fait de donner à l'élément i un poids proche de $\bar{p}_j(i)$, respectivement. Si pour une requête j , l'algorithme de recherche retourne un élément i au rang $P_j(i)$, alors la concordance entre le classement de l'algorithme et celui des utilisateurs est mesurée par le pourcentage du classement selon la similarité $S_j(i)$

$$S_j(i) = \sum_{k=P_j(i)-\frac{\sigma_j(i)}{2}}^{P_j(i)+\frac{\sigma_j(i)}{2}} Q_i(i, k)$$

Le tracé de $S_j(i)$ comme fonction de $P_j(i)$ décrit la performance de recherche d'un algorithme. Un $S_j(i)$ élevé indique un niveau de précision élevé de l'algorithme.

1.4.4 Moyenne du rang de recherche normalisé

Le *average of normalized modified retrieval rank* (ANMRR) est une mesure objective utilisée par MPEG-7 pour évaluer les performances de recherche [36]. Cette mesure combine la précision et le rappel en une seule valeur objective. Pour l'utiliser, des requêtes et des ensembles dits *ground truth* (vérités de terrain) sont choisis manuellement. Pour chaque requête, un ensemble de d'éléments *ground truth* qui sont les plus pertinents pour la requête seront identifiés. Ces éléments ne seront ordonnés d'aucune manière qui soit. Un bon descripteur est sensé retourner tous les éléments *ground truth* pour une requête image.

Soit $NG(q)$ le nombre d'éléments *ground truth* pour la requête q , en premier lieu $K(q)$ éléments seront retournés pour la requête q par le système, un $K(q)$ acceptable est déterminé par $K(q) = \min(4 \cdot NG(q), 2 \cdot GTM)$ ou GTM est le maximum des $NG(q)$ pour toutes les requêtes de la base de données. À chaque image $item_k$ qui est retrouvée parmi les premiers $K(q)$ éléments, une valeur de rang $Rank(item_k)$ est assignée. Cette valeur sera égale à l'ordre selon lequel est retrouvée l'image. Ainsi le rang de la première image sortie est 1 et celui de la dernière est $K(q)$. On assignera le rang $(K(q) + 1)$ à toutes les autres images qui n'ont pas été retourné par le système parmi les premiers $K(q)$. Le NMRR sera donc donné par :

$$NMRR(q) = \frac{\sum_{k=1}^{NG(q)} \frac{Rank(item_k)}{NG(q)} - 0.5 - \frac{NG(q)}{2}}{K(q) + 0.5 - 0.5 \cdot NG(q)} \quad (1.4)$$

Le $NMRR$ est dans l'intervalle $[0,1]$, plus cette mesure est petite meilleure est la performance de la recherche. Prenons un exemple pour illustrer cette mesure : pour la requête q , les éléments *ground truth* seront $I_{q1}, I_{q2}, \dots, I_{q10}$, ce qui veut dire que $NG(q) = 10$. Le résultat idéal dans ce cas sera que les 10 premières recherches soient exactement les éléments

$I_{q1}, I_{q2}, \dots, I_{q10}$, leurs rangs respectifs seront $Rank(I_{q1})=1, Rank(I_{q2})=2, \dots, Rank(I_{q10})=10$.

Le numérateur dans l'équation (eq1.4) sera : $\sum_{h=1}^{10} \frac{Rank(I_{qh})}{10} - 0.5 - \frac{10}{2} = 0$.

ANMRR sera finalement défini comme la moyenne des NMRR sur le nombre de requêtes NQ et sera donné par :

$$ANMRR = \frac{1}{NQ} \sum_{q=1}^{NQ} NMRR(q)$$

1.5 Conclusion

Nous avons décrit dans ce chapitre les différentes étapes de la description de forme. Après avoir présenté les différentes classifications des descripteurs et quelques uns des plus connus d'entre eux, nous avons passé en revue certaines mesures de similarité et les moyens d'évaluer un descripteur. Il apparaît cependant qu'il y a matière à choisir d'autres propriétés de l'image lors de sa description. Et c'est à ce propos que nous introduisons le prochain chapitre.

Chapitre 2

Topologie, homologie et théorie de Morse

2.1 Introduction

La majorité des disciplines scientifiques qui génèrent des simulations computationnelles font usage de procédés d'enregistrement et de capteurs pour produire des données image, il paraîtra donc trivial à chacun que ceci produit une quantité gigantesque de données à traiter en plus du fait que la complexité de ces données croît avec la sophistication des capteurs. Tandis que le fait d'avoir un si grand volume d'informations nous donne une plus grande confiance dans la capture des propriétés étudiées, extraire celles qui sont les plus pertinentes s'avère très compliqué. Les techniques automatisées d'analyse de ces données sont en général divisées en deux catégories : de haut et de bas niveau.

Les exemples de méthodes de haut niveau d'analyse sont la reconnaissance d'objets, la reconnaissance de caractères et d'écritures manuscrites, ce qui induit bien sûr la description formelle des caractéristiques de l'objet. Tandis que les méthodes de bas niveau s'intéressent plus aux structures géométriques des objets imagés et sont donc utilisées comme première étape ou prétraitement par des méthodes de haut niveau. Et c'est dans cette optique que l'homologie computationnelle a été prouvée comme pouvant servir de base pour des analyses poussées.

Que mesure la topologie?

Considérons la Figure 2.1:



Figure 2.1 – (a) une forme consistant en une seule composante connexe ; (b) forme divisée en deux composantes connexes

En topologie la propriété la plus importante qui différencie ces deux formes n'est pas qu'elles soient de tailles inégales ni d'apparences différentes, mais plutôt que dans Figure 2.1(a) l'objet est formé d'une composante connexe alors que dans Figure 2.1(b) il est constitué de deux composantes connexes. La topologie est donc une science qui étudie les propriétés géométriques invariantes d'un objet quand celui-ci est étiré, tordu ou rétréci de manière continue [37].

2.2 Topologie algébrique

Nous présentons, dans cette section, une brève revue des notions basiques et de la terminologie de la topologie algébrique. Pour une description plus complète, nous referons le lecteur à n'importe quel texte standard dans le domaine, tels que [38, 39, 40, 41, 42, 43], ou un texte plus récent [37] qui utilise un ordinateur pour développer une approche de calcul combinatoire pour le sujet. Les définitions suivantes sont simplifiées et restreignent leur domaine à des sous-ensembles de \mathbb{R}^n avec une topologie de métrique Euclidienne.

2.2.1 Homéomorphismes et variétés

Etant donné que la topologie s'intéresse à la connexité de la forme, elle s'appuie sur la capacité à produire des équivalences entre formes ayant la même connexité. Cette équivalence est définie par l'homéomorphisme.

Définition 2.1 Soient les deux espaces topologiques X et Y et soit la fonction $f: X \rightarrow Y$ qui est continue et bijective. Et soit sa fonction inverse $f^{-1}: Y \rightarrow X$ définie comme $f^{-1}(y) = \{x: f(x) = y\}$, continue elle aussi. X est alors homéomorphe à Y et la transformation f est dite *homéomorphisme*.

En imagerie, on représente communément les formes par leurs surfaces, appelées variétés/tubulures. La topologie utilise cette notion d'homéomorphisme pour définir une variété de dimension n , comme étant un ensemble qui est localement homéomorphe à un sous-ensemble d'un hyperplan de dimension n , comme dans la figure 2.2.

Définition 2.2 Un ensemble $M \subset \mathbb{R}^m$ est une variété de dimension n si et seulement si chaque point $p \in M$ a un voisinage homéomorphe avec un sous-ensemble ouvert de \mathbb{R}^n . En plus un ensemble $M \subset \mathbb{R}^m$ est une variété de dimension n dite avec frontière si et seulement si chaque point $p \in M$ a un voisinage homéomorphe avec soit un sous-ensemble ouvert de \mathbb{R}^n ou alors de $\mathbb{R}^{n-1} \times [0, \infty)$.

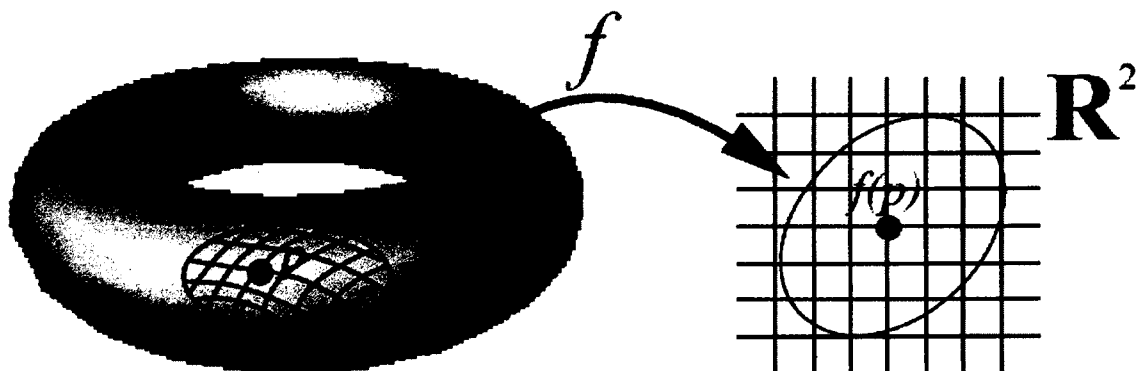


Figure 2.2 - L'homéomorphisme f transforme un voisinage de p sur le tore en un voisinage de l'origine du plan. Une telle transformée existe pour tout point du tore, sachant que le tore est une variété de dimension 2.

2.2.2 Equivalence homotopique

Une autre méthode pour indiquer que deux formes possèdent la même connexité est l'équivalence homotopique. En imagerie, une homotopie est une interpolation continue entre deux courbes, deux surfaces ou deux images.

Définition 2.3 Si f_0 et f_1 sont des fonctions continues de l'espace X vers l'espace Y . On dit que f_0 est homotopique à f_1 si et seulement s'il existe une fonction continue $F : X \times [0, 1] \rightarrow Y$ telle que $F(x, 0) = f_0(x)$ et $F(x, 1) = f_1(x)$ pour tout x . La transformée F est appelée *homotopie* entre f_0 et f_1 .

Définition 2.4 Une fonction continue $f : X \rightarrow Y$ est appelée *équivalence homotopique* s'il existe une fonction $g : Y \rightarrow X$ telle que $g \circ f$ est homotopique à la fonction identité dans X et $f \circ g$ est homotopique à la fonction identité dans Y . La transformée g est considérée comme une homotopie inverse pour la transformée f . Deux ensembles équivalents en homotopie seront considérés comme ayant le même *type d'homotopie*.

Ainsi, deux ensembles qui ont la même connexité sont dits équivalents en homotopie, et si deux formes sont équivalentes en homotopie elles vont posséder le même nombre de trous et de composants.

Mais l'homotopie reste généralement difficile à calculer algorithmiquement, c'est pourquoi, une méthode plus simple pour démontrer l'équivalence homotopique, utilise une transformation spéciale appelée *déformation en rétraction*.

Définition 2.5 Soit $A \subset X$. Une transformée continue $r : X \rightarrow A$ est dite rétraction si et seulement si $r(a) = a$ pour tout $a \in A$, un tel ensemble A est appelé *retract* de X . S'il existe une fonction $H : X \times [0,1] \rightarrow X$ telle que pour tout $x \in X$, $H(x,0) = x$ et $H(x,1) = r(x)$, alors H est appelée une déformation rétraction et A un retract de déformation de X . L'ensemble X a le même type homotopique que chacun de ses retracts de déformation $A \subset X$. L'équivalence homotopique est plus faible que l'équivalence homeomorphique, par exemple le cercle (variété de dimension 1) et le tore (variété de dimension 3) sont homotopiques (le cercle est un retract du tore) mais pas homéomorphes.

Exemple 2.1 Soit X l'anneau décrit dans la Figure 2.3. On le paramétrise avec des coordonnées polaires (r, θ) , $1 \leq r \leq 2$, $0 \leq \theta \leq 2\pi$. Soit A le cercle unitaire et f l'application qui va de X vers A donnée par $f(r, \theta) = (1, \theta)$. Alors, f est une équivalence homotopique, car l'inclusion du cercle dans X fournit l'application inverse homotopique g . Il est facile de voir que f est un retract de X en A et l'application $h(r, \theta, t) = tf(r, \theta) + (1-t)(r, \theta)$ définit une déformation en rétraction de l'anneau en un cercle unitaire.

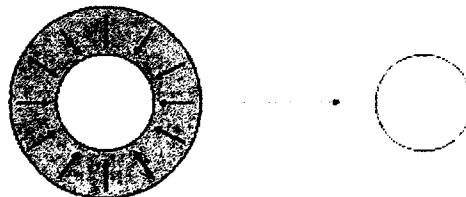


Figure 2.3 - Le cercle est un retract de déformation de l'anneau

2.3 Homologie

Dans la plus part des cas il s'avère difficile de déterminer si deux transformations entre espaces sont homotopiques ou que deux espaces sont homotopiquement équivalents. En pratique, lorsqu'on veut comparer des espaces ou des transformations à l'aide d'outils topologiques, on peut le faire d'une manière très satisfaisante, en comparant leurs structures d'homologie respectives, ce qui est moins général mais plus facile à calculer. Même si l'homologie est moins intuitive que l'homotopie, sa nature combinatoire la rend plus attractive. Son grand avantage est qu'elle peut être utilisée pour dire quand est-ce que deux espaces ne sont pas équivalents en homotopie ou homéomorphes, sans avoir besoin de prouver qu'il n'existe aucun homéomorphisme ou aucune équivalence homotopique entre les deux.

En fait, l'homologie se propose d'attribuer à tout objet géométrique X une suite de groupes abéliens¹ $(H_n(X))_{n \geq 0}$ [44]. Intuitivement parlant, l'homologie d'un espace topologique est un invariant qui compte le nombre de trous dans cet espace. Pour des sous-ensembles d'espaces tridimensionnels, trois groupes d'homologie non triviaux H_0 , H_1 et H_2 existent, leurs rangs respectifs β_0 , β_1 et β_2 (appelés *nombre de Betti*) déterminent respectivement le nombre de composantes connexes, le nombre de boucles (tunnels) indépendantes et le nombre de chambres vides (de cavités) indépendantes dans le sous-ensemble. Le fait que l'homologie intègre ces propriétés en fait un invariant topologique tout à fait utile pour décrire une forme.

Pour des complexes cellulaires, la construction des groupes d'homologie est basée sur le concept de frontière des cellules. Intuitivement, la frontière d'un point est l'ensemble vide qui correspond au 0 algébrique, la frontière d'une cellule unidimensionnelle est formée de ses deux points extrêmes. De façon similaire, la frontière d'une cellule à deux dimensions sera son contour. Un raisonnement analogue nous permettra de déterminer les frontières de

¹ Un groupe abélien est une structure algébrique dont les éléments peuvent être additionnés, soustraits, multipliés par des scalaires entiers, mais la division par des scalaires n'est pas permise.

cellules de plus grandes dimensions. Lorsque ces cellules sont le résultat d'une triangularisation de l'espace, elles seront appelées *simplexes* et sont assemblées suivant certaines règles. Elles formeront ainsi un *complexe simplicial*.

Définition 2.6 Un complexe simplicial X est un ensemble de simplexes de toutes dimensions organisés de telle manière que l'intersection de deux d'entre eux est soit vide soit un simplexe du même complexe.

La triangularisation en imagerie et dans les systèmes dynamiques mène naturellement à la construction de blocs cubiques qui peuvent être utilisés de la même manière que des simplexes pour définir une homologie appelée *homologie cubique*. La structure cubique nous permet de simplifier les définitions de la théorie de l'homologie et rend plus facile l'utilisation de l'analyse topologique des images, et ce pour deux raisons. La première est mathématique et consiste dans le fait que le produit cartésien de deux simplexes n'est pas nécessairement un simplexe, alors que le produit de cubes est toujours un cube. La deuxième est informatique et consiste dans le fait qu'une image est une collection de pixels qui sont représentés par des carrées, ce qui rend l'image elle-même un complexe cubique.

Nous nous proposons maintenant de donner quelques bases pour le calcul de l'homologie.

2.4 Homologie cubique

Définition 2.7 Un intervalle élémentaire est un intervalle fermé $I \subset \mathbb{R}$ de la forme

$$I = [l, l+1] \text{ ou } I = [l, l] = \{l\} \quad l \in \mathbb{N} ;$$

Lorsque l'intervalle élémentaire consiste en un seul point, il est dit dégénéré, alors que ceux de longueur 1 sont non-dégénérés. Ensuite, à partir des intervalles élémentaires, on peut construire les cubes élémentaires.

Définition 2.8 Un cube élémentaire $\sigma \subset \mathbb{R}^n$ est un produit fini d'intervalles élémentaires,

$$\sigma = I_1 \times I_2 \times \dots \times I_n$$

Comme énoncé par les définitions qui suivent, chaque cube élémentaire possède une dimension, une ou plusieurs faces et possiblement une ou plusieurs cofaces, autres que lui-même.

Définition 2.9 La dimension de σ est définie comme le nombre d'intervalles élémentaires non-dégénérés dans σ , dénoté par le nombre $q \in \{0, 1, \dots, n\}$.

Notons que le préfixe q -, ou bien q disposé en indice, sera souvent utilisé pour désigner un objet de dimension q . Ainsi, un q -cube élémentaire dénote un σ_q , soit un cube élémentaire de dimension q .

Pour $q \geq 1$, soit $J = \{k_0, k_1, \dots, k_{q-1}\}$ le sous-ensemble ordonné de $\{1, 2, \dots, n\}$ des indices pour lesquels $I_{k_j} = [a_j, b_j]$ est non-dégénéré ; On définit

$$A_{k_j} \sigma = I_1 \times \dots \times I_{k_j-1} \times \{a_j\} \times I_{k_j+1} \times \dots \times I_n$$

Et

$$B_{k_j} \sigma = I_1 \times \dots \times I_{k_j-1} \times \{b_j\} \times I_{k_j+1} \times \dots \times I_n$$

Il est facilement vérifiable que chacune de ces i -faces est un $(q-1)$ -cube élémentaire. Elles sont ainsi dites $(q-1)$ -face de σ . De la même manière on pourra définir les $(q-2)$ -faces, ... jusqu'à la 0-face de σ .

Définition 2.10 La frontière du q -cube élémentaire σ consiste en la somme alternée de ses $(q-1)$ -faces i.e.

$$\partial_q \sigma = \sum_{i=0}^{q-1} (-1)^i (B_{k_i} \sigma - A_{k_i} \sigma)$$

Par exemple dans Figure 2.4, $\partial_2 A = b + g - f - a$, $\partial_1 b = 4 - 3$ et $\partial_0 3 = 0$.

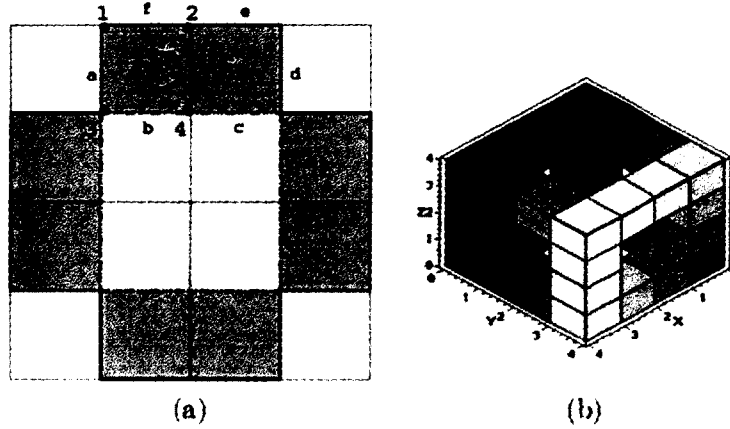


Figure 2.4 - Exemple de complexes, (a) bidimensionnel et (b) tridimensionnel. Dans (a) le complexe est formé par les carrés grisés.

Définition 2.11 Soient σ^1 et σ^2 deux cubes élémentaires tels que $\sigma^1 \subseteq \sigma^2$, alors σ^1 est une face de σ^2 et inversement σ^2 est une coface de σ^1 . Lorsque σ^1 est une face de σ^2 et que $\dim \sigma^1 < \dim \sigma^2$ alors σ^1 est une face propre de σ^2 et inversement σ^2 est une coface propre de σ^1 . Puis, si σ^1 est une face propre de σ^2 et que $\dim \sigma^1 = \dim \sigma^2 - 1$, alors σ^1 est une face primaire de σ^2 et inversement σ^2 est une coface primaire de σ^1 [45].

Dans Figure 2.4, on peut voir un 2-cube élémentaire A, avec ses 1-faces a, b, g, f. Les 0-faces de A sont 1, 2, 3 et 4.

Définition 2.12 Un ensemble $X \subset \mathbb{R}^d$ est cubique si X peut être formé d'une union finie de cubes élémentaires.

À partir des cubes élémentaires, on peut former une structure appelée complexe cubique.

Définition 2.13 Un complexe cubique $\mathcal{K} \subset \mathbb{R}^n$ est une collection de q -cubes élémentaires où $0 \leq q \leq n$ et qui respecte :

1. Toute face d'un cube élémentaire de \mathcal{K} est aussi contenue dans \mathcal{K} .
2. L'intersection de deux cubes élémentaires de \mathcal{K} est soit vide, soit une face partagée par les deux.

Notons que le complexe associé à un ensemble topologique cubique X sera dénoté par $\mathcal{K}(X)$. La dimension m de \mathcal{K} est par définition le plus grand nombre q pour lequel \mathcal{K} contient un q -cube élémentaire. Finalement, à partir de la structure d'un complexe, il est possible de construire ce qu'on appelle les chaînes cubiques. Soit un m -complexe cubique \mathcal{K} , on dénote par $E_q = \{\sigma_q^1, \dots, \sigma_q^{N_q}\}$, l'ensemble de tous les q -cubes élémentaires de \mathcal{K} .

Définition 2.14 Une q -chaîne c est une somme formelle des éléments de E_q , soit

$$c = \sum_{i=1}^{N_q} \lambda_i \sigma_q^i \quad \text{où } \lambda_1, \dots, \lambda_{N_q} \in \mathbb{Z}.$$

Pour chaque $q \geq 0$, on associe au complexe \mathcal{K} , le groupe libre $C_q(\mathcal{K})$, généré par l'ensemble des q -chaînes. Deux q -chaînes sont sommées en faisant la somme de leurs coefficients correspondants. Etant donné que l'on ne s'intéresse qu'aux complexes finis, les groupes $C_q(\mathcal{K})$ sont générés de manière finie et $C_q(\mathcal{K}) = 0$ si $q > m$ et naturellement $C_q(\mathcal{K}) = 0$ si $q < 0$.

Définition 2.15 (l'opérateur de bord) La frontière algébrique d'une q -chaîne c d'un complexe \mathcal{K} correspond à sa frontière géométrique où chaque face est visitée selon un certain ordre. Cet ordre implique l'existence d'une orientation du complexe qui consiste en un classement linéaire des sommets du complexe. Cela revient donc à faire la somme de ses $(q-1)$ -faces. Le bord noté ∂ de l'ensemble $E_q = \{\sigma_1, \dots, \sigma_n\}$ du complexe \mathcal{K} est donc une

extension linéaire de la définition du bord d'un q -cube élémentaire dans la définition 2.10. On obtient alors la projection $\partial_q : C_q(\mathcal{K}) \rightarrow C_{q-1}(\mathcal{K})$.

Notons que $\partial_0 = 0$ puisque $\partial_{-1}(\mathcal{K}) = 0$. Cette frontière satisfait aussi à la propriété très importante $\partial^2 = 0$ i.e. $\partial_q \circ \partial_{q+1} = 0$ pour tout q , voir [39] pour les preuves. L'information algébrique extraite de la structure cubique donnée peut être exprimée comme un complexe de chaîne fini et libre (C, ∂) qui est une séquence de groupes abéliens connectés par les frontières, donné par

$$0 \xrightarrow{\partial_{n+1}} C_n(\mathcal{K}) \xrightarrow{\partial_n} C_{n-1}(\mathcal{K}) \xrightarrow{\partial_{n-1}} C_{n-2}(\mathcal{K}) \xrightarrow{\partial_{n-2}} \dots C_0(\mathcal{K}) \xrightarrow{\partial_0} 0$$

Dans le sillage de la définition de l'opérateur de bord on définit les deux groupes suivants :

$$Z_k(\mathcal{K}) := \ker(\partial : C_k \rightarrow C_{k-1}) = \{c \in C_k : \partial c = 0\}$$

qui est le groupe des k -cycles de \mathcal{K} , et

$$B_k(\mathcal{K}) = \text{Im}(\partial : C_k \rightarrow C_{k-1}) = \{c \in C_k : \exists c' \in C_{k+1}, c = \partial c'\}$$

qui est le groupe des k -bords de \mathcal{K}

On a :

$$B_k(\mathcal{K}) \subset Z_k(\mathcal{K}) \subset C_k(\mathcal{K})$$

Enfin, à chaque cycle $c \in Z_k(\mathcal{K})$ est associée sa classe d'homologie $c + B_k(\mathcal{K}) = \{c + b : b \in B_k(\mathcal{K})\}$. Deux cycles c et c' seront donc homologues si et seulement si $\exists b \in B_k(\mathcal{K}) : c' = c + b$.

2.5 Les groupes d'homologie

Tel que présentée précédemment, l'homologie assigne à tout espace topologique X une séquence de groupes abéliens $(H_n(X))_{n \geq 0}$ possédant les propriétés suivantes :

1. Fonctorialité. Pour toute fonction $f: X \rightarrow Y$ est associée une séquence d'homomorphismes induits $f_*: H_n(X) \rightarrow H_n(Y)$, $n \geq 0$, tels que $(g \circ f)_* = g_* \circ f_*$ et $(i_X)_* = i_{H_n(X)}$ où i_X est la fonction identité dans X . Notons que certains texts standards utilisent la notation $H_n(f)$ au lieu de f_* .

2. Invariance homotopique. Si $f, g: X \rightarrow Y$ sont homotopiques, alors $f_* = g_*$.

Le k^{ieme} groupe d'homologie du complexe \mathcal{K} décrivant l'objet X est donné par le groupe quotient:

$$H_k(\mathcal{K}) = Z_k(\mathcal{K}) / B_k(\mathcal{K})$$

Les nombres de Betti de \mathcal{K} seront ensuite donnés par : $\beta_k(\mathcal{K}) = \text{Rang}(H_k(\mathcal{K}))$.

Lemme de l'invariance topologique: Si \mathcal{K} et \mathcal{K}' sont deux complexes dont les supports sont homéomorphes, alors leurs groupes d'homologie sont isomorphes et leurs nombres de Betti sont égaux. Ce résultat reste vrai si les supports de \mathcal{K} et \mathcal{K}' ont le même type d'homotopie.

Considérons, pour poursuivre notre illustration, la Figure 2.5 et notons par X_i les objets décrits dans les figures Figure 2.5(i) avec $i = a, b, c, d, e$. La Figure 2.5(a) est composée d'un point libre et d'un segment de droite, aucun trou n'apparaît ni aucune cavité, ce qui implique les groupes d'homologie suivants : $H_0(X_a) \cong \mathbb{Z}^2$ (2 composantes connexes), $H_1(X_a) \cong 0$ (nombre de tunnels indépendants), $H_2(X_a) \cong 0$ (nombre de cavités). Et les nombres de Betti seront $(\beta_0(X_a), \beta_1(X_a), \beta_2(X_a)) = (2, 0, 0)$

$H_0(X_b) \cong \mathbb{Z}$, $H_1(X_b) \cong \mathbb{Z}$ (1 tunnel), $H_2(X_b) \cong 0$. Avec $(\beta_0(X_b), \beta_1(X_b), \beta_2(X_b)) = (1, 1, 0)$

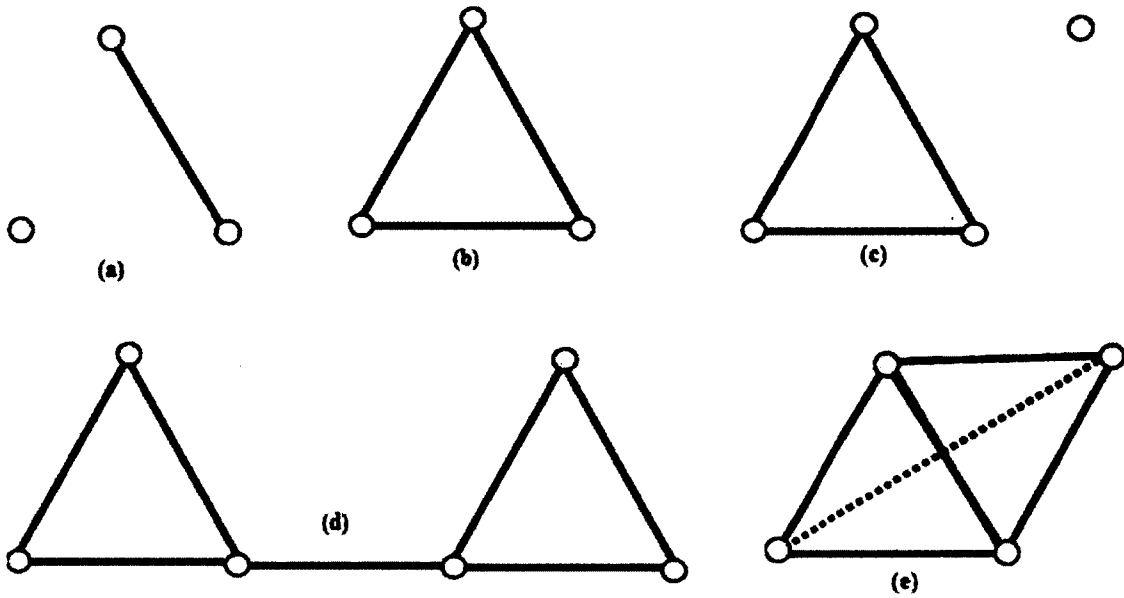


Figure 2.5 – Différents scénarios pour illustrer le calcul des groupes d'homologie

$H_0(X_c) \cong \mathbb{Z}^2$, $H_1(X_c) \cong 0$ (si on inclut l'intérieur du triangle dans l'objet X_c , le tunnel disparaît), $H_2(X_b) \cong 0$. Avec $(\beta_0(X_c), \beta_1(X_c), \beta_2(X_c)) = (2, 0, 0)$.

$H_0(X_d) \cong \mathbb{Z}$, $H_1(X_d) \cong \mathbb{Z}^2$ (Apparition d'un deuxième tunnel), $H_2(X_d) \cong 0$. Avec $(\beta_0(X_d), \beta_1(X_d), \beta_2(X_d)) = (1, 2, 0)$.

$H_0(X_e) \cong \mathbb{Z}$, $H_1(X_e) \cong 0$, $H_2(X_e) \cong \mathbb{Z}$ (ici l'objet inclut les faces grisées mais est vide à l'intérieur ce qui crée une cavité). Avec $(\beta_0(X_e), \beta_1(X_e), \beta_2(X_e)) = (1, 0, 1)$.

2.6 Calcul d'homologie

2.6.1 Méthode classique

La méthode classique de calcul d'homologie et des nombres de Betti passe par le calcul de matrices dites d'incidence qui expriment l'opérateur de bord par rapport aux bases canoniques. Cette méthode est illustrée dans l'exemple suivant :

Soit la forme dans le plan suivant :

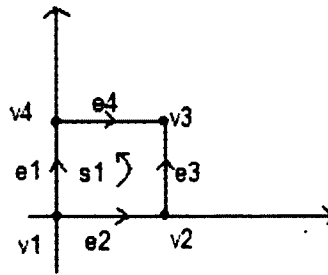


Figure 2.6

Il est évident que $\partial_0 = 0$ (puisque $C_{-1} = 0$) et donc $Z_0 = C_0$. On aura ainsi :

$$\dots \rightarrow 0 \xrightarrow{\partial_1} C_2 \simeq \mathbb{Z} \xrightarrow{\partial_2} C_1 \simeq \mathbb{Z}^4 \xrightarrow{\partial_1} C_0 \simeq \mathbb{Z}^4 \xrightarrow{\partial_0} 0 \dots$$

∂_1 peut s'écrire sous la forme d'une matrice en mettant les arêtes en colonnes et les sommets en lignes, ce qui donne :

$$\begin{matrix} & e1 & e2 & e3 & e4 \\ \begin{matrix} v1 \\ v2 \\ v3 \\ v4 \end{matrix} & \begin{pmatrix} -1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & -1 \end{pmatrix} \end{matrix}$$

Pareil pour ∂_2 , ou on met les arêtes en lignes et l'unique surface en colonne:

$$\begin{matrix} & s1 \\ e1 & \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \\ e2 & \\ e3 & \\ e4 & \end{matrix}$$

Notons que le signe des éléments de la matrice est positif ou négatif selon que l'orientation de l'arête soit compatible à celle de l'objet ou pas. On calcule ensuite les groupes Z_k, B_k avec

$$k = 0, 1 : \quad \left. \begin{array}{l} \text{Im } \partial_2 = B_1 = \mathbb{Z} \\ \text{ker } \partial_1 = Z_1 = \mathbb{Z} \end{array} \right\} \Rightarrow H_1 = \frac{\mathbb{Z}}{\mathbb{Z}} = 0$$

$$\text{Et} \quad \left. \begin{array}{l} \text{Im } \partial_1 = B_0 = \mathbb{Z}^3 \\ \text{ker } \partial_0 = Z_0 = \mathbb{Z}^4 \end{array} \right\} \Rightarrow H_0 = \frac{\mathbb{Z}^4}{\mathbb{Z}^3} = \mathbb{Z}$$

Notre complexe est donc équivalent à un carré plein avec une composante connexe et 0 tunnels indépendants. Rappelons que le Kernel (ou noyau) noté \ker d'une matrice A consiste en l'ensemble de tous les vecteurs $x = (x_1, x_2, \dots, x_n)$ pour qui $Ax = 0$ avec $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$ et $x \in \mathbb{R}^m$. L'image de la matrice A , quant à elle, notée Im est l'ensemble de tous les vecteurs de \mathbb{R}^n qui sont égaux à Ax . Autrement dit $\text{Im } A$ est l'ensemble des vecteurs $y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ pour lesquels le système d'équations $Ax = y$ admet une solution.

2.6.2 Méthode de calcul par la réduction

2.6.2.1 Effondrement (écroulement) élémentaire

Le lecteur aura sûrement remarqué que même un ensemble cubique très simple peut contenir un grand nombre de cubes élémentaires, ce qui alourdira considérablement le temps de calcul surtout lors des opérations sur les matrices et des groupes Z_k et B_k . Nous présenterons donc ici une méthode qui vise à réduire le nombre de cubes élémentaires servant à calculer l'homologie du groupe. Pour ce, affinons un peu plus le concept de faces propres (cf.

définition 2.11) mais cette fois-ci dans le contexte d'ensembles cubiques au lieu de cubes élémentaires.

Définition 2.16 Soit l'ensemble cubique X et $\mathcal{K}(X)$ le complexe cubique qui lui est associé. Soit aussi $Q \in \mathcal{K}(X)$, si Q n'est face propre d'aucun $P \in \mathcal{K}(X)$, alors Q est dite face maximale dans X . $\mathcal{K}_{\max}(X)$ sera l'ensemble des faces maximales dans X . Une face qui est face propre d'exactly un seul cube dans X sera dite face libre dans X .

Définition 2.17 Soit Q une face propre dans X et soit P l'unique cube dans $\mathcal{K}(X)$ tel que Q est une face propre de P . Soit $\mathcal{K}'(X) := \mathcal{K}(X) \setminus \{Q, P\}$. On définit

$$X' := \bigcup_{R \in \mathcal{K}'(X)} R$$

Alors X' est un espace cubique obtenu à partir de X via un effondrement élémentaire de P par Q . En plus, on peut prouver que $\mathcal{K}(X') = \mathcal{K}'(X)$.

Exemple 2.2 Soit $X = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ (Figure 2.7 [37]). Alors

$$\begin{aligned}\mathcal{K}_2(X) &= \{[0, 1] \times [0, 1]\}, \\ \mathcal{K}_1(X) &= \{[0] \times [0, 1], [1] \times [0, 1], [0, 1] \times [0], [0, 1] \times [1]\}, \\ \mathcal{K}_0(X) &= \{[0] \times [0], [0] \times [1], [1] \times [0], [1] \times [1]\}.\end{aligned}$$

Il y a quatre faces libres dans X i.e. les éléments de $\mathcal{K}_1(X)$. Soit $Q = [0, 1] \times [1]$, alors Q est face propre de $P = [0, 1] \times [0, 1]$.

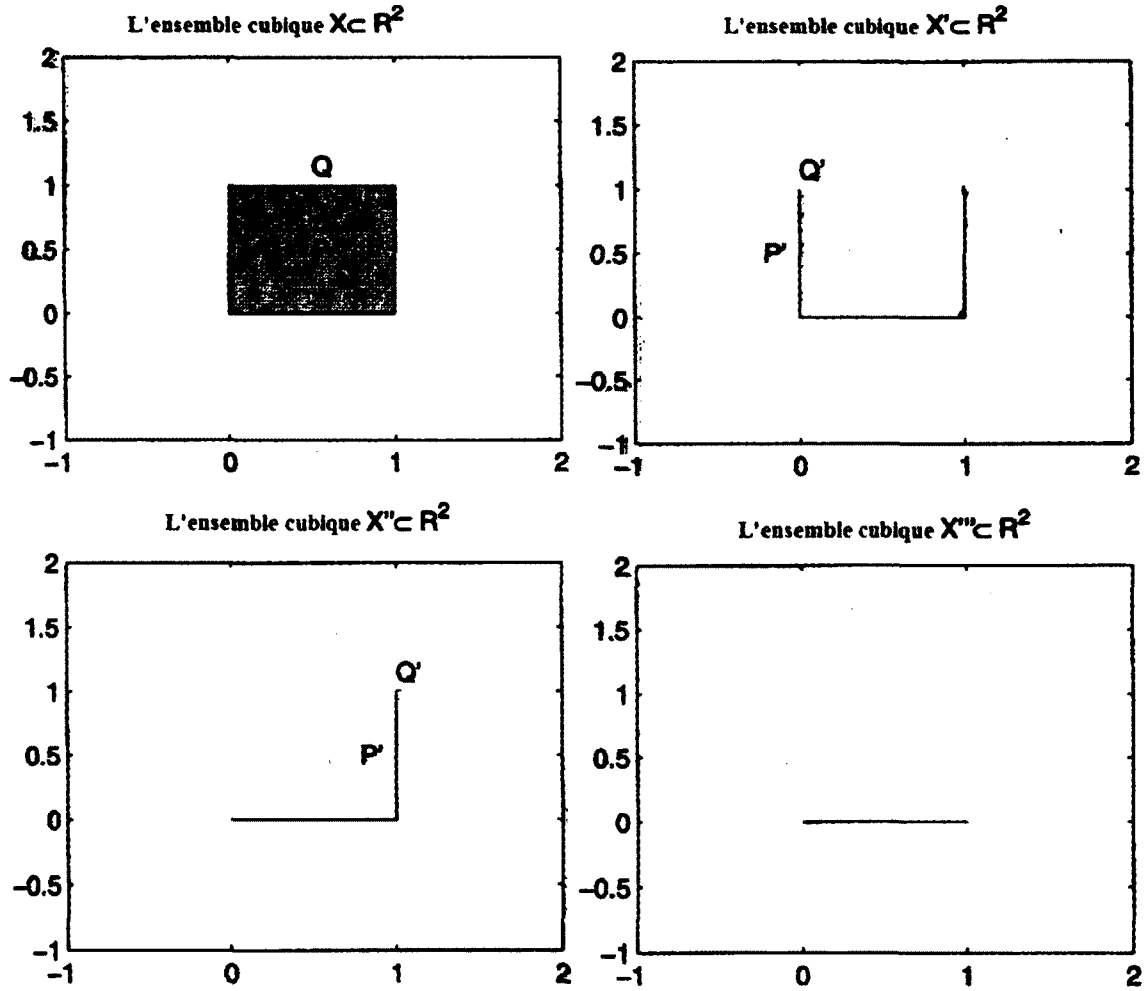


Figure 2.7 – Séquence d'effondrements élémentaires de $[0,1] \times [0,1] \subset \mathbb{R}^2$

Si on pose X' comme l'espace cubique obtenu à partir de X par l'effondrement de P par Q , alors $X' = [0] \times [0,1] \cup [1] \times [0,1] \cup [0,1] \times [0]$ et :

$$\mathcal{K}_1(X') = \{[0] \times [0,1], [1] \times [0,1], [0,1] \times [0]\}$$

$$\mathcal{K}_0(X') = \{[0] \times [0], [0] \times [1], [1] \times [0], [1] \times [1]\}$$

Observons que les faces libres de X' sont différentes de celles de X . En particulier, $[0] \times [1]$ et $[1] \times [1]$ qui sont des faces libres avec $[0] \times [1]$ face propre de $[0] \times [0, 1]$. X'' sera l'espace obtenu par l'effondrement de $[0] \times [0, 1]$ par $[0] \times [1]$ et :

$$\begin{aligned}\mathcal{K}_1(X'') &= \{[1] \times [0, 1], [0, 1] \times [0]\} \\ \mathcal{K}_0(X'') &= \{[0] \times [0], [1] \times [0], [1] \times [1]\}\end{aligned}$$

Même chose pour X'' , où l'effondrement de $[1] \times [0, 1]$ par $[1] \times [1]$ donnera X''' pour qui :

$$\begin{aligned}\mathcal{K}_1(X''') &= \{[0, 1] \times [0]\} \\ \mathcal{K}_0(X''') &= \{[0] \times [0], [1] \times [0]\}\end{aligned}$$

Un effondrement final de $[0, 1] \times [0]$ par $[1] \times [0]$ résultera en un point unique $X'''' = [0] \times [0]$. Ainsi, et à travers ce processus nous aurons réduit un 2-cube contenant 9 cubes élémentaires en un complexe cubique consistant en un sommet unique. Il a été prouvé, par ailleurs, que deux complexes cubiques reliés par un effondrement élémentaire auront la même homologie.

Théorème 2.1 Soit X un ensemble cubique et X' obtenu à partir de X via un effondrement de $P_0 \in \mathcal{K}_k(X)$ par $Q_0 \in \mathcal{K}_{k-1}(X)$. Alors

$$H_*(X') \cong H_*(X).$$

Corollaire Soit $Y \subset X$ deux ensembles cubiques. En supposant que Y peut être obtenu à partir de X via une série d'effondrements élémentaires. Alors

$$H_*(Y) \cong H_*(X).$$

2.6.2.2 Méthode de réduction

Cette méthode généralise l'idée de l'effondrement élémentaire des groupes élémentaires aux groupes de chaînes élémentaires [46-50]. Ainsi on désire maintenant réduire la taille du

complexe de chaîne utilisé pour calculer l'homologie, ce qui résulte en l'algorithme *CCR* « *Chain Complex Reduction* »[37] dont on présentera une version simplifiée. Le principe algébrique est le suivant :

Considérons le complexe de chaînes libre (C, ∂) donné par

$$0 \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_{n+2}} C_{n+1} \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_k} C_{k-1} \xrightarrow{\partial_{k-1}} \dots C_0 \xrightarrow{\partial_0} 0$$

Pour chaque k , on associe une base E_k pour C_k (formée uniquement de k -chaînes élémentaires) on définit ensuite la forme bilinéaire suivante

$\langle ., . \rangle : C_k \times C_k \rightarrow \mathbb{Z}$ sur les générateurs $u, v \in E_k$ par :

$$\langle u, v \rangle = \begin{cases} 1 & \text{si } u = v \\ 0 & \text{sinon.} \end{cases}$$

Soit $m \in \{1, \dots, n\}$, $A \in E_m$ et $a \in E_{m-1}$ tels que a soit une face propre de A , on pourra donc écrire

$$\partial_m A = \lambda a + r, \quad \text{avec } \lambda = \pm 1$$

On effectue par la suite une réduction locale du complexe, ce qui revient à :

- Supprimer A de E_m et a de E_{m-1} de telle façon à préserver l'homologie de (C, ∂) .
- Définir un complexe de chaînes réduit $(\bar{C}, \bar{\partial})$ avec une nouvelle base

$$\bar{E}_k := \begin{cases} E_k & \text{si } k \neq m-1, m \\ E_{m-1} \setminus \{a\} & \text{si } k = m-1 \\ E_m \setminus \{A\} & \text{si } k = m \end{cases}$$

et l'opérateur de frontière réduit $\bar{\partial}$

$$\bar{\partial}_k v := \begin{cases} \partial_k v & \text{si } k \neq m, m+1 \\ \partial_k v - \lambda^{-1} \langle \partial_k v, a \rangle \partial_k A & \text{si } k = m \\ \partial_k v - \langle \partial_k v, A \rangle A & \text{si } k = m+1 \end{cases}$$

Théorème 2.2 $(\bar{C}, \bar{\partial})$ est un complexe de chaînes libres et $H(C) = H(\bar{C})$.

Dans Figure 2.8, on présente une version plus formelle de l'algorithme [47].

```

Input: Base  $E_2$  de 2-cellules
Output: #trous et composantes connexes

begin

  Crée bases  $E_0$  et  $E_1$ ;
  while  $E_2$  non vide
    choisir  $A$  dans  $E_2$  et une face propre  $a$  dans  $E_1$ ;
     $E_2 := E_2 \setminus \{A\}$ ;
     $E_1 := E_1 \setminus \{a\}$ ;

  endwhile;

  while  $E_1 \neq \emptyset$  et  $\exists A \in E_1$  et  $\exists a \in E_0$  t.q  $a$  est face
  de  $A$ .
    REDUCE( $A, a$ );

  endwhile;

end;
```

Figure 2.8 – Algorithme HOMOLOGY

Avant de définir la procédure REDUCE(), on note par $T(B)$ l'ensemble de toutes les faces propres pour $\forall B \in E_1$, et on note $R(a)$ l'ensemble de toutes les cofrontières de a , $\forall a \in E_0$.

```

Procedure REDUCE( $A, a$ );

  for each  $B$  in  $R(a) \setminus \{A\}$  do
     $T(B) := [T(B) \cup T(A)] \setminus \{a\}$ ;
  endfor;

  for each  $b$  in  $T(A) \setminus \{a\}$  do
     $R(b) := [R(b) \cup R(a)] \setminus \{A\}$ ;
  endfor;

   $E_1 := E_1 \setminus \{A\}$ ;
   $E_0 := E_0 \setminus \{a\}$ ;

end REDUCE;

```

Figure 2.9 – Procédure REDUCE()

Pour un complexe planaire, cet algorithme est linéaire avec le nombre de générateurs. Pour un complexe de dimension m (avec $m \geq 3$), la complexité devient d'environ $O(n^3)$ où n est la cardinalité maximale des bases de générateurs de toutes les dimensions supérieures à m . Cependant, avec un choix judicieux des générateurs, l'on peut diminuer cette complexité à $O(n^2 \log n)$.

2.7 Homologie relative

On peut étendre les définitions précédentes de l'homologie à une paire d'espaces (X, A) avec $A \subseteq X$. L'idée est d'ignorer tout générateur dans A . Plus précisément, les groupes d'homologie relative sont définis tels qui suit. Soit $C_n(X, A)$ le groupe quotient $C_n(X)/C_n(A)$. Il s'en suit que toutes les n -chaines dans A seront triviales dans $C_n(X, A)$. Etant donné que la fonction frontière $\partial: C_n(X) \rightarrow C_{n-1}(X)$ va de $C_n(A)$ à $C_{n-1}(A)$, elle induit une fonction de frontière quotient $\partial: C_n(X, A) \rightarrow C_{n-1}(X, A)$. En variant n , on obtient une séquence de fonctions frontière

$$\dots \rightarrow C_n(X, A) \xrightarrow{\partial} C_{n-1}(X, A) \rightarrow \dots$$

La relation $\partial \circ \partial = 0$ reste valable pour ces fonctions frontiere vu que elle l'était déjà avant de passer aux groupes quotients. On a donc un complexe de chaînes et les groupes d'homologie $\ker \partial / \text{Im } \partial$ de ce complexe de chaînes seront par définition les *groupes d'homologie relative* $H_n(X, A)$. A partir de la définition de la fonction de frontiere relative, on peut voir que :

1. Les groupes d'homologie $H_n(X, A)$ sont représentés par des cycles relatifs : les n -chaînes $\alpha \in C_n(X)$ telles que $\partial \alpha \in C_{n-1}(A)$.
2. Un cycle relatif α est trivial dans $H_n(X, A)$ si et seulement si il est une frontiere triviale : $\alpha = \partial \beta + \gamma$ pour un certain $\beta \in C_{n+1}(X)$ et $\gamma \in C_n(A)$.

Ces propriétés rendent plus précise l'idée intuitive selon laquelle $H_n(X, A)$ serait l'homologie de X modulo A et que l'on illustre dans l'exemple suivant.

Exemple 2.3 Soit X le complexe de la Figure 2.10 (a) dont l'espace sous jacent est l'intervalle et soit A le sous-complexe consistant en la frontiere de cet intervalle. Il est, dès lors, facile de voir que la 1-chaîne $c = e$ est un 1-cycle relatif de X modulo A et que tout 1-cycle de X modulo A est forcément un multiple de c . Etant donné qu'il n'y a pas de frontieres en dimension 1, il s'en suit que $H_1(X, A) \cong \mathbb{Z}$. Et vu que toute 0-chaîne dans X est aussi dans A , on aura donc $H_0(X, A) = 0$. Considérons à présent le complexe X dans la Figure 2.10(b) consistant en l'intervalle et le point et soit A la frontiere de l'intervalle. Pour des raisons similaires à celles dans (a), on aura $H_1(X, A) \cong \mathbb{Z}$. On peut observer que les 0-chaînes basées sur les sommets V_0 et V_1 sont dans A et ne mènent à aucun générateur dans l'homologie relative. les 0-chaînes multiples de V_2 ne sont pas dans A et ne sont frontieres d'aucune 1-chaîne. Il s'en suit qu'à partir de la définition de l'homologie relative, V_2 est générateur d'une classe d'homologie relative. Et ainsi $H_0(X, A) \cong \mathbb{Z}$.

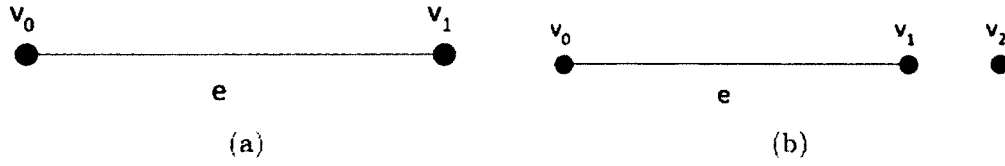


Figure 2.10 - L'intuition derrière l'homologie relative

Il arrive parfois qu'en ignorant la structure de A on obtient de l'information sur X . Ainsi, et tel qu'illustré dans l'exemple précédent, le rang du groupe d'homologie relative de niveau zéro $H_0(X, A)$ donne le nombre de composantes dans X qui n'intersectent pas avec A . On obtient des groupes d'homologie réduite en réduisant H_0 d'un générateur. Ces groupes sont dénotés $H_i^\#$ et ils sont reliés aux groupe d'homologie relative comme ceci : $H_i^\#(X) \cong H_i(X, \{p\})$ où p est n'importe quel point de X . Pour certaines paires d'espaces (X, A) (voir [43], proposition 2.22), les groupes d'homologie réduite permettent de donner une meilleure interprétation des groupes d'homologie relative. On rappelle qu'étant donné un espace topologique X et un sous-espace A de X , on peut identifier tous les points dans A à une classe d'équivalence unique et laisser les points en dehors de A équivalents à eux même. L'espace quotient résultant dénoté X/A , dans lequel A sera identifié à un seul point. Pour certaines paires (X, A) , il est prouvé dans [43] que

$$H_n(X, A) = H_n(X/A, [A]) = H_n^\#(X/A), \text{ pour tout } n.$$

Les groupes d'homologie sont calculés en utilisant les algorithmes dans [47,48,49,50]. De légères modifications aux algorithmes d'homologie dans ces références nous permettent de calculer d'une manière très efficace les groupes d'homologie relative. L'homologie calculée avec des coefficients entiers peut produire des coefficients de torsion tel que l'on peut voir à partir du théorème fondamental de groupe abéliens générés de manière finie (voir [41] (Théorème 4.3)).

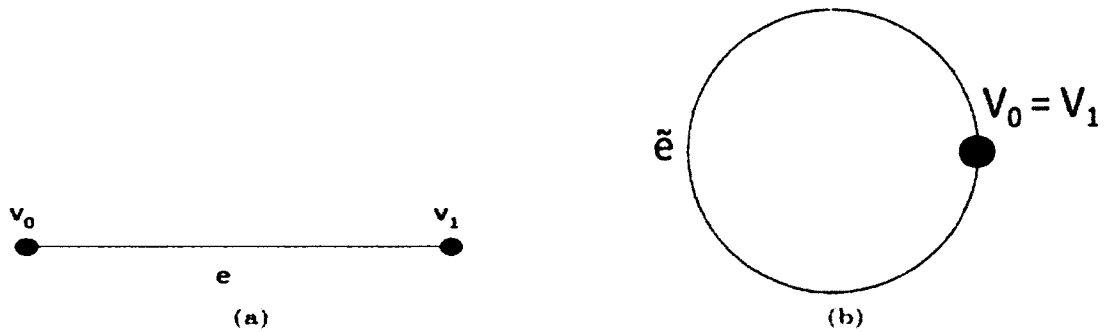


Figure 2.11 - Les groupes d'homologie dans le complexe (a) modulo les sommets V_0 et V_1 sont isomorphes aux groupes d'homologie réduite de l'espace quotient dans (b) duquel on peut directement voir que $H_0(X, A) = 0$ et $H_1(X, A) = \mathbb{Z}$

Les groupes d'homologie calculés avec des coefficients appartenant à un intervalle ne produisent, par contre, pas de coefficients de torsion et sont en fait des espaces vectoriels. Etant donné que les objets en 2D et 3D n'ont pas de torsion, il suffit d'utiliser les groupes d'homologie sur un domaine. Dans ce qui suit nous utiliserons les groupes d'homologie avec des coefficients réels.

2.8 La suite exacte d'homologie d'une paire

Tel que mentionné précédemment, une des utilités de l'homologie relative est de donner de l'information au sujet des groupes d'homologie absolus. La meilleure façon pour formuler la relation entre les groupe d'homologie relative $H_n(X, A)$ et les groupes absolus $H_n(X)$ et $H_n(X)$ sera à travers le concept de *suite exacte d'homologie d'une paire*, concept introduit par Eilenberg.

Définition 2.18 Une suite de groupes et homomorphismes

$$\dots \rightarrow G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3 \dots$$

est dite être **exacte à G_2** si $\text{image } f_1 = \text{kernel } f_2$

Si elle est exacte partout, elle sera dite **suite exacte**. Une **longue suite exacte** est une suite exacte dont l'ensemble des indices est l'ensemble des entiers.

Théorème 2.3 Soit le complexe X et A un sous-complexe. Il existe alors une longue suite exacte

$$\dots \rightarrow H_p(A) \xrightarrow{i_*} H_p(X) \xrightarrow{\pi_*} H_p(X, A) \xrightarrow{\partial_*} H_{p-1}(A) \dots$$

où $i: A \rightarrow X$ et $\pi: (X, \emptyset) \rightarrow (X, A)$ sont des inclusions et ∂_* est induite par l'opérateur de frontière ∂ et appelée *homéomorphisme de frontière d'homologie*. La preuve de ce théorème est de nature algébrique et peut être trouvée dans [41].

Proposition 2.1 Soit A un retract de déformation de X . Alors, le groupe d'homologie relative $H_p(X, A) = 0$ pour tout $p \in \mathbb{N}$.

Preuve. Supposons que A soit une retract de déformation de X . On aura donc à partir de la définition de la déformation-rétraction, on déduit tout de suite que l'inclusion $i: A \rightarrow X$ induit un isomorphisme $i_*: H_*(A) \rightarrow H_*(X)$ en homologie vu que $r \circ i$ est égal à la fonction identité sur A , et $i \circ r$ est homotopique à l'identité de X .

À partir de la séquence exacte de la paire (X, A) , il s'en suit que pour tout p

$$\ker \pi_* = \operatorname{Im} i_* = H_p(X)$$

Ainsi π_* disparaît sur $H_p(X)$. D'un autre côté, vu que

$$\ker \partial_* = \operatorname{Im} \pi_* = 0 \text{ et } \operatorname{Im} \partial_* = \ker i_* = 0$$

Il s'ensuit immédiatement que ∂_* est injective et disparaît sur $H_p(X, A)$, pour tout $p \in \mathbb{N}$.

On a donc $H_p(X, A) = 0$ pour tout $p \in \mathbb{N}$.

2.9 La séquence exacte d'homologie d'un triplet

Une généralisation facile de la longue séquence exacte d'une paire donnera la longue séquence exacte d'un triplet qui s'avère très utile car elle établira des propriétés de notre descripteur de forme (chapitre 3).

Définition 2.19 Etant donné le triplet d'espaces (X, A, B) avec $X \supset A \supset B$, on aura une courte séquence exacte de complexes de chaînes

$$0 \rightarrow \frac{C(A; \mathbb{R})}{C(B; \mathbb{R})} \xrightarrow{i} \frac{C(X; \mathbb{R})}{C(B; \mathbb{R})} \xrightarrow{j} \frac{C(X; \mathbb{R})}{C(A; \mathbb{R})} \rightarrow 0$$

où $i: (A, B) \subset (X, B)$ et $j: (X, B) \subset (X, A)$ sont des inclusions canoniques. Ceci nous amène à la *séquence exacte d'homologie d'un triplet*

$$\dots \xrightarrow{\partial_*} H_k(A, B; \mathbb{R}) \xrightarrow{i_*} H_k(X, B; \mathbb{R}) \xrightarrow{j_*} H_k(X, A; \mathbb{R}) \xrightarrow{\partial_*} H_{k-1}(A, B; \mathbb{R}) \xrightarrow{i_*} \dots$$

où i_* et j_* sont respectivement induits par les inclusions canoniques i, j et ∂_* est l'homomorphisme de connexion (voir [41,42] pour plus de détails sur ces constructions).

2.10 Théorie de Morse en imagerie

La théorie de Morse est une partie de l'algèbre topologique. Elle s'intéresse aux effets des points critiques sur la topologie d'un ensemble. C'est pour cela que nous nous proposons de présenter, très rapidement, des notions d'algèbre topologique qui nous semblent utiles aux lecteurs pour la compréhension de la suite.

2.10.1 Points critiques et fonctions de Morse

Soit f une fonction réelle et différentiable sur la variété M . Un point $p \in M$ est critique si et seulement si ses dérivées par rapport à un système de coordonnées locales sur M sont nulles ou tendent vers zéro.

Plus spécifiquement, étant donné que M est une variété lisse de dimension n , alors pour tout point $p \in M$. Il existe une correspondance point-à-point différentiable g entre un voisinage dans M et un sous-ensemble ouvert de \mathbb{R}^n telle que : $g(p) = x = (x_1, x_2, \dots, x_n)$. Le point $p \in M$ sera critique dans f si le gradient :

$$\nabla f = \left(\frac{\partial f \circ g^{-1}(x)}{\partial x_1}, \frac{\partial f \circ g^{-1}(x)}{\partial x_2}, \dots, \frac{\partial f \circ g^{-1}(x)}{\partial x_n} \right) = 0$$

La théorie de Morse s'intéresse plus particulièrement aux points critiques non-dégénérés. De tels points, appelés aussi « *points de Morse* », sont des points critiques pour qui le Hessien :

$$V(f) = J(\nabla f) = \begin{bmatrix} \frac{\partial \nabla f}{\partial x_1} \\ \frac{\partial \nabla f}{\partial x_2} \\ \vdots \\ \frac{\partial \nabla f}{\partial x_n} \end{bmatrix}$$

Ce Hessien, disons nous, aura un déterminant non-nul. Et vu que $\partial^2 f / \partial x_i \partial x_j = \partial^2 f / \partial x_j \partial x_i$, la matrice sera symétrique. Soient $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ les valeurs propres de $V(f)$, si une de ces valeurs propres est nulle, le point critique est dégénéré sinon il sera appelé *non-dégénéré*.

Définition 2.20 Une fonction lisse $f : M \rightarrow \mathbb{R}$ est une fonction de Morse si tous ses points critiques, sont non-dégénérés, c'est à dire que près d'un point critique p , il existe un système de coordonnées locales (x_1, x_1, \dots, x_n) , dans lequel la matrice du Hessien de f en p est non-singulière (c.-à-d. possède un inverse)[45,51]. Les Figure 2.12 et Figure 2.13 illustrent quelques exemples de points critiques dégénérés et non-dégénérés.

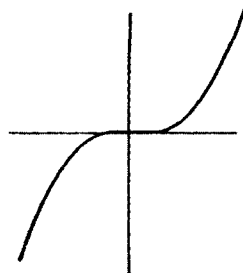


Figure 2.12 - $\varphi(x) = x^3$ a un point critique dégénéré en 0.

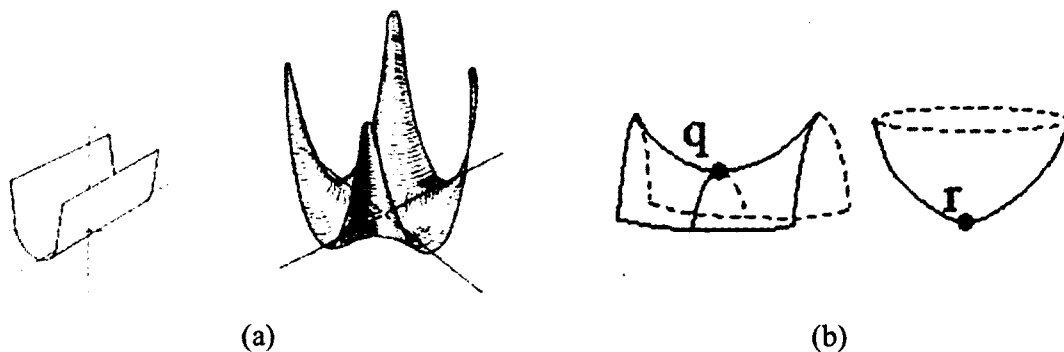


Figure 2.13 - Exemples (a) : points critiques dégénérés, (b) : points critiques non- dégénérés

À chaque point critique p non-dégénéré est associé un invariant, que Morse a appelé l'indice de f en p , et dénoté $\lambda(p)$. Cet indice est défini comme le nombre de valeurs propres négatives du Hessien $V(f)$ et correspond à la dimension de la variété instable $W^u(p) = \left\{ q \in M \mid \lim_{t \rightarrow -\infty} \psi(t, q) = p \right\}$ où $\psi : \mathbb{R} \times M \rightarrow M$ dénote le flot du gradient négatif de f .

Intuitivement parlant, l'indice de Morse d'un point critique p est le nombre de directions indépendantes autour de p dans lesquelles f diminue. La Figure 2.14 illustre cette notion d'indice d'un point critique selon la dimension de la variété instable $W^u(p)$.

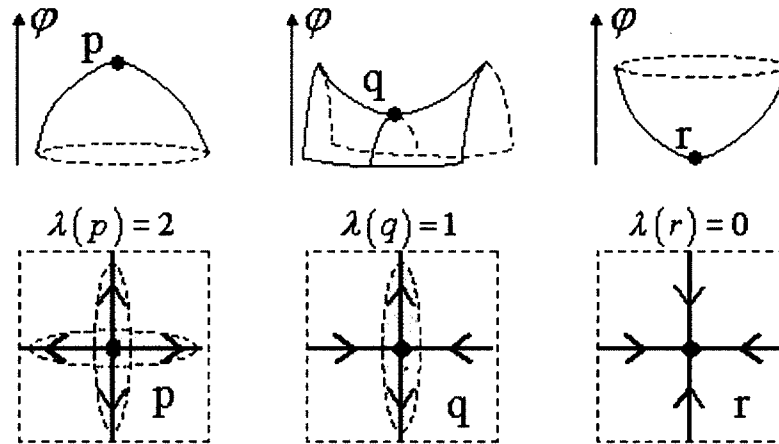


Figure 2.14 - Exemple de l'indice d'un point critique selon la dimension de sa variété instable

2.10.2 Lemme de Morse

Le lemme de Morse est une affirmation très puissante sur le comportement des fonctions autour de points critiques non-dégénérés. Il stipule que le voisinage d'un point critique non-dégénéré peut être déformé en un voisinage d'un point critique non-dégénéré d'une fonction quadratique cette fois-ci. En d'autres termes, la première fonction agira comme une fonction quadratique dans ce voisinage, peu importe sa véritable classe, et le type de cette fonction quadratique sera déterminé par l'index du point critique en question.

Lemme Soit p un point critique non-dégénéré de f avec λ comme index, il existera alors un système de coordonnées locales $(y_1(x), y_2(x), \dots, y_n(x))$ dans un voisinage U de p avec p comme origine et

$$f(x) = f(p) - y_1(x)^2 - y_2(x)^2 - \dots - y_\lambda(x)^2 + y_{\lambda+1}(x)^2 + \dots + y_n(x)^2$$

Pour la preuve de ce lemme voir [51]. La théorie de Morse s'intéresse donc à déterminer le type d'homotopie d'une forme en se basant sur ses points critiques. Un exemple classique [52] démontre les effets des points critiques sur l'homotopie, en observant la portion d'un tore en dessous d'un plan de coupe pendant que ce plan bouge à travers le tore vers le haut. Un effet similaire peut être observé en plongeant un beigne dans une tasse de café comme dans les figures Figure 2.15 et Figure 2.16.

Soit M une variété de dimension 2 qui décrit la surface du tore et soit $f(p)$ la fonction qui retourne la hauteur du point $p \in M$. En posant le bas du tore comme la hauteur 0 et le sommet comme hauteur 1. Notons M^a les points $p \in M$ tels que $f(p) \leq a$, ces points sont la partie du tore de hauteur inférieur ou égale à a . Au fur et à mesure que le plan de coupe se déplace vers le haut, la figure 2.16 montre les changements dans la topologie du tore. Ces changements peuvent être décrits par le processus d'attachement d'une k -cellule appropriée à la surface tronquée et ce à chaque étape (cf. section 2.10.3). On remarque que la dimension de la cellule attachée est égale à l'indice du point critique qui vient d'être traversé par le plan de coupe.

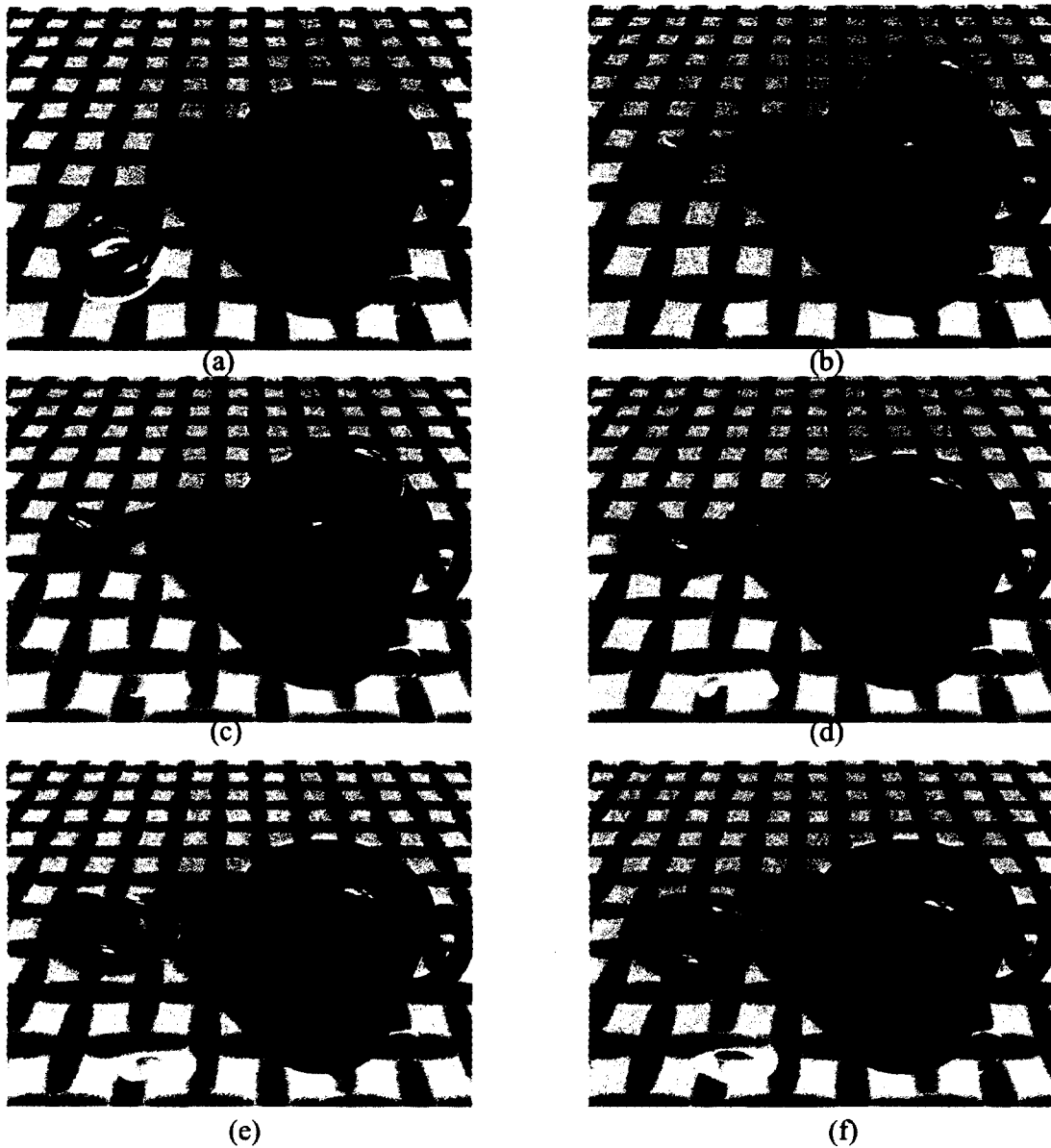


Figure 2.15 - Plonger un beigne dans une tasse de café. (b) dès que le beigne touche la surface du café, la portion de sa surface immergée change de l'ensemble vide à une forme homéomorphe au disque. (c) à l'instant où le café atteint le trou du beigne, la topologie de la portion immergée change d'un disque à un cylindre tronqué comme dans (d). À l'instant où le trou est immergé en entier comme dans (e), la topologie de la portion immergée change du cylindre tronqué à un tore tronqué comme dans (f). En relâchant le beigne, le café recouvrira tout le beigne, ce qui complètera la forme du tore

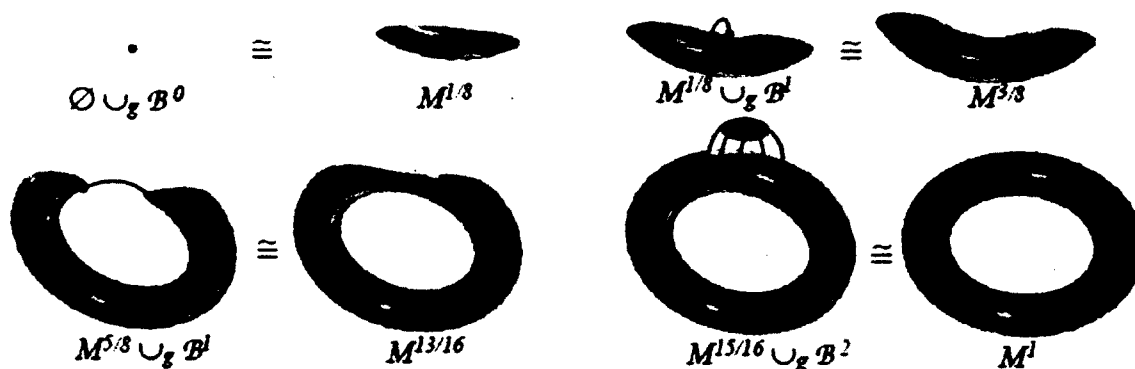


Figure 2.16 - les différentes classes d'homotopie du tore coupé

2.10.3 Attacher une k -cellule

Souvent, on peut « altérer » topologiquement un ensemble pour avoir la même classe d'homotopie qu'un autre ensemble. L'équivalence en homotopie est basée sur le nombre de composants et de trous dans l'ensemble. De fait, une disjonction ou une fusion va changer l'homotopie de l'ensemble en modifiant le nombre de composants. Une autre opération peut, elle aussi, changer le genre de l'ensemble, c'est l'attachement d'une k -cellule.

Un ensemble dans un espace euclidien est dit *fermé* s'il contient sa frontière. La *fermeture* de l'ensemble est l'union de cet ensemble avec sa frontière. Un ensemble est *borné* s'il est possible de l'entourer par une certaine sphère de rayon fini centrée à son origine (celle de l'ensemble). Ces deux notions combinent dans l'espace euclidien pour former la notion de compacité.

Définition 2.21 Un ensemble $A \subset \mathbb{R}^n$ est compact s'il est borné et fermé.

La boule fermée et la sphère sont des représentations typiques d'ensembles dans des espaces données.

Définition 2.22 Une k -boule (fermée) est un ensemble de vecteurs k -dimensionnels de longueurs inférieures ou égales à 1.

$$B^k = \{x \in \mathbb{R}^k : \|x\| \leq 1\}.$$

Définition 2.23 Une k -sphère S^k est un ensemble de vecteurs de dimension $(k+1)$ et de longueurs égales à 1.

$$S^k = \{x \in \mathbb{R}^{k+1} : \|x\| = 1\}$$

La frontière d'une k -boule dans \mathbb{R}^k est une $(k-1)$ -sphère. Une 0 -boule est un point unique et sa frontière (dans $\mathbb{R}^0 = \{0\}$) est l'ensemble vide. On désigne communément la frontière d'une k -cellule e^k par ∂e^k .

Enfin, si deux ensembles ne partagent pas la même classe d'équivalence homotopique, souvent l'un pourra être converti en l'autre à travers un processus d'attachement d'une ou plusieurs k -cellules.

Définition 2.24 Le processus d'attachement d'une k -cellule à un ensemble Y avec la transformation $g : S^k \rightarrow Y$ est dénoté :

$$Y \cup_g B^k$$

et est obtenu en effectuant la somme topologique (union disjointe) de Y et B^k , ensuite on identifie chaque $x \in S^{k-1}$ avec $g(x) \in Y$.

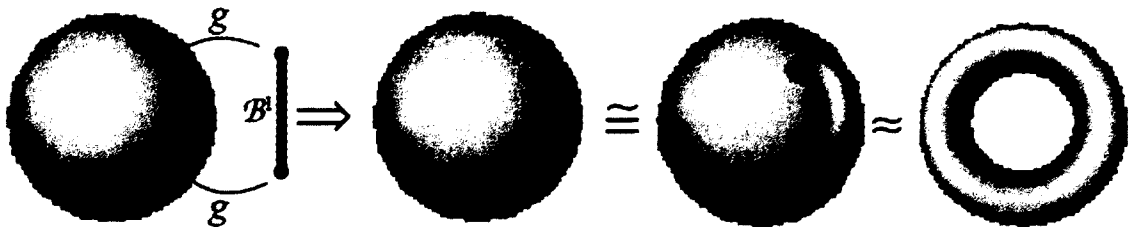


Figure 2.17 - L'attachement d'une 1-cellule à une 3-boule conduit à une forme homotopique au tore.

La Figure 2.17 décrit le processus d'attachement d'une k -cellule ; L'ensemble Y initial est dans ce cas une 3-boule. On attache une 1-cellule en utilisant la fonction g pour identifier les deux points qui forment la frontière de sa 0-sphère (à la 1-cellule) avec les points sur la 2-sphère qui est la surface de la 3-boule. Cette 1-cellule attachée est une déformation en rétraction de sa version plus épaisse, et la 3-boule avec cette version épaissie, attachées, sont homéomorphe au tore.

2.10.4 Théorèmes de Morse

Théorème 2.4 [51] Soit $f: M \rightarrow \mathbb{R}$ une C^2 , soit $a < b$ et supposons que l'ensemble $f^{-1}[a, b]$ est compact et ne contient aucun point critique de f , alors M^a est homéomorphe à M^b .

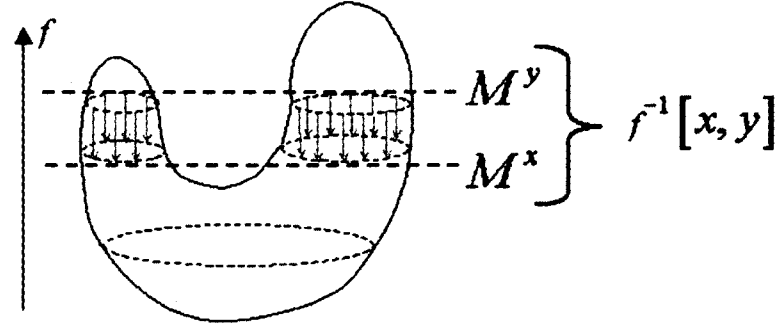
Ce théorème montre qu'un sous-ensemble d'une variété M^a est topologiquement similaire à un sous-ensemble plus petit $M^b \subset M^a$, s'il n'apparaît aucun point critique dans M^a qui n'est pas déjà dans M^b (voir Figure 2.18).

Théorème 2.5 Soit $f: M \rightarrow \mathbb{R}$ une C^2 et soit $p \in M$ un point critique non-dégénéré avec λ comme indice. En posant $f(p) = c$ et en supposant que $f^{-1}[c - \varepsilon, c + \varepsilon]$ est compact et ne contient aucun point critique autre que p pour un $\varepsilon > 0$ donné. Alors pour tout ε suffisamment petit, l'ensemble $M^{c+\varepsilon}$ a le même type d'homotopie que $M^{c-\varepsilon}$ avec une λ -cellule attachée (Figure 2.19) :

$$M^{c+\varepsilon} \cong M^{c-\varepsilon} \cup_{\varphi_p} e_p^\lambda$$

$\varphi_p: \partial e_p^\lambda \rightarrow M^{c-\varepsilon}$ étant la fonction d'attachement de la cellule e_p^λ , elle consiste en une application continue, qui attache la λ -cellule e_p^λ à $M^{c-\varepsilon}$ par le biais de sa frontière. Le nouvel espace résultant de l'opération d'attachement $M^{c-\varepsilon} \cup_{\varphi_p} e_p^\lambda$, correspond à l'union disjointe de $M^{c-\varepsilon}$ et e_p^λ quotientée par la relation d'équivalence qui identifie chaque point

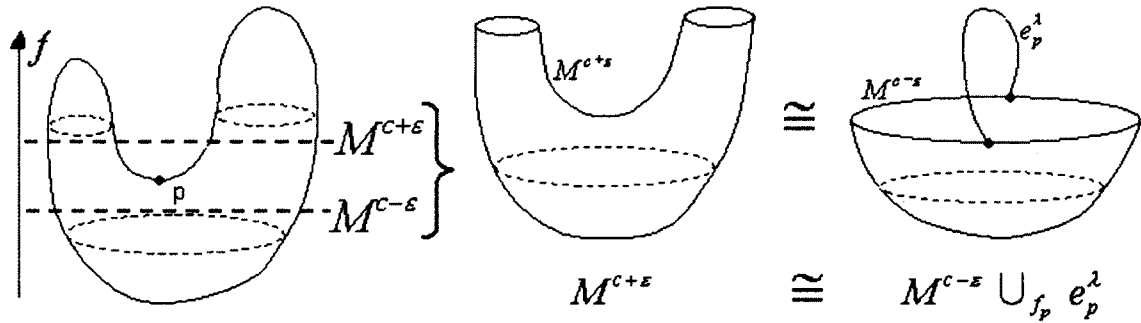
$s \in \partial e_p^\lambda$ avec $\varphi_p(s) \in M^{c-\varepsilon}$. Notons que la fonction d'attachement φ_p doit être définie pour tout $x \in \partial e_p^\lambda$, c'est-à-dire que la frontière de e_p^λ au complet doit être attachée à $M^{c-\varepsilon}$ [45].



$f^{-1}([x, y])$ ne contient aucun point critique, alors $M^y \cong M^x$ et

M^x est un rétract de déformation de M^y .

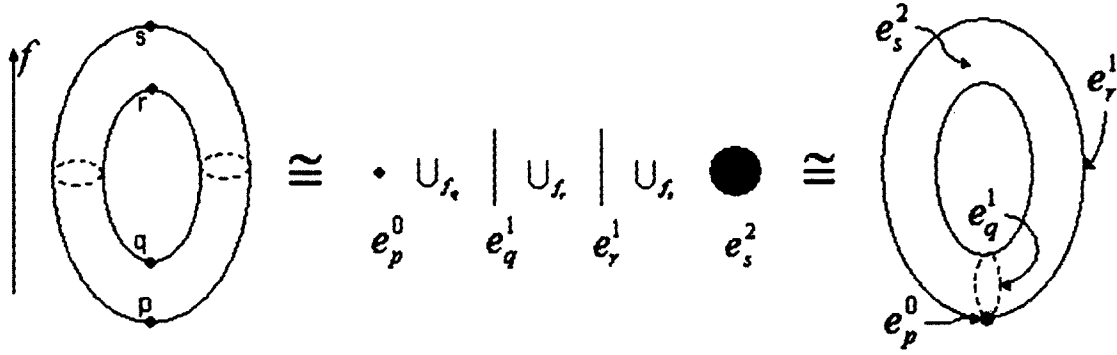
Figure 2.18 - Exemple du théorème 2.4.



Contient un point critique d'indice , est homotopiquement équivalent à une cellule attachée à .

Figure 2.19 – Exemple du théorème 2.5

En décomposant un tore en une union de cellules élémentaires, comme dans la Figure 2.20, on obtient une 0-cellule, deux 1-cellules et une 2-cellule. Sachant que les groupes d'homologie des niveaux 0, 1 et 2 du tore sont respectivement isomorphes à \mathbb{Z}^1 , \mathbb{Z}^2 et \mathbb{Z}^1 , ceci soulève la question de savoir s'il existe un lien entre la décomposition en cellules élémentaires et les groupes d'homologie de M . Le théorème 2.6[51] permet d'établir ce lien, où l'attachement d'une λ -cellule affecte les générateurs d'homologie réduite. Rappelons que l'homologie réduite, dénotée $H_k^\#$, est obtenue en réduisant H_0 d'un générateur.



Dans cet exemple, le tore se décompose en l'union de quatre cellules élémentaires, soit une 0-cellule, deux 1-cellules et une 2-cellule.

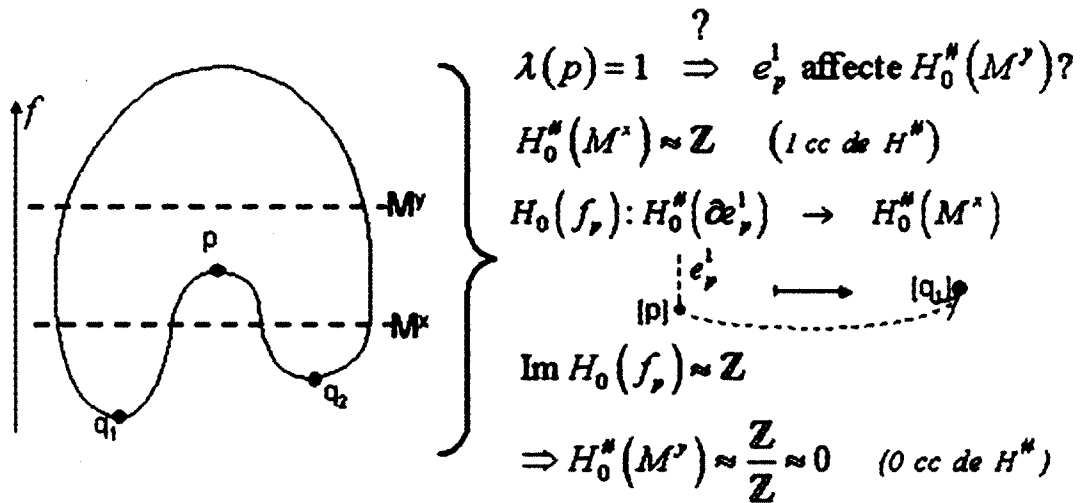
Figure 2.20 - Exemple de la décomposition d'une variété en cellules élémentaires

Théorème 2.6 Soient $x, y \in \mathbb{R}$ tels que $x < y$ et supposons que $M^y \cong M^x \cup_\varphi e^\lambda$, où φ est la fonction d'attachement de la λ -cellule e^λ à l'ensemble M^x . Alors,

1. Si $k \neq \lambda, k \neq \lambda - 1$, alors $H_k^\#(M^y) \cong H_k^\#(M^x)$.
2. $H_{\lambda-1}^\#(M^y) \cong \frac{H_{\lambda-1}^\#(M^y)}{\text{Im } H_{\lambda-1}(\varphi)}$.
3. $H_\lambda^\#(M^y) \cong H_\lambda^\#(M^x) \oplus \ker H_{\lambda-1}(\varphi)$.

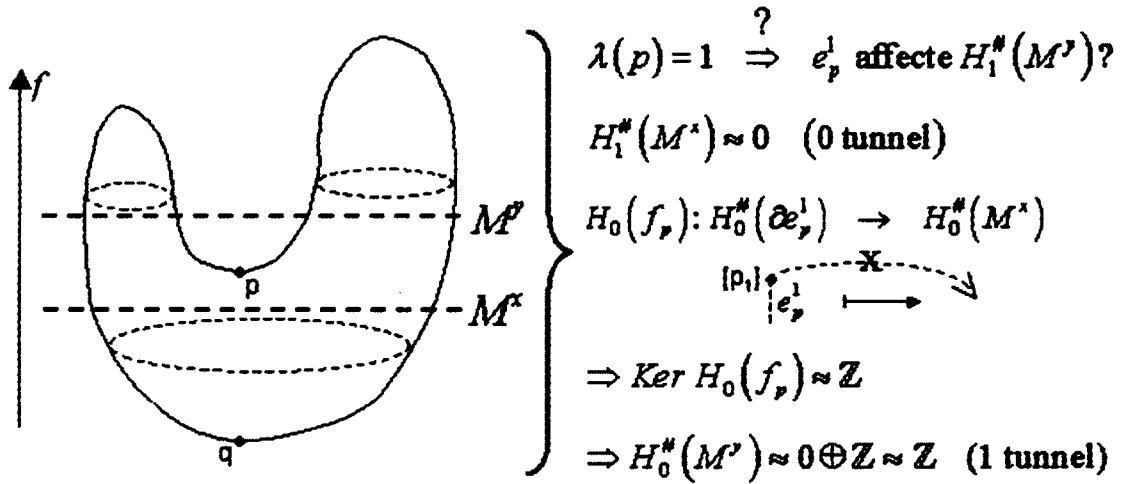
Ce théorème signifie que selon la nature de la fonction d'attachement φ , l'ajout de la λ -cellule e^λ peut seulement affecter l'homologie des niveaux λ et $\lambda-1$. Quand c'est $H_{\lambda-1}^\#$ qui est affecté, la λ -cellule rend équivalents deux générateurs du niveau $\lambda-1$ et réduit le nombre de Betti $\beta_{\lambda-1}$ de un. Voir la Figure 2.21 pour un exemple.

Lorsque que c'est plutôt $H_\lambda^\#$ qui est affecté, la λ -cellule crée un nouveau générateur au niveau λ et par conséquent incrémente le nombre de Betti β_λ de un. Voir la Figure 2.22 pour un tel exemple.



$[q_1]$ et $[q_2]$ sont rendus équivalents sur le plan de l'homologie par l'ajout de e_p^1 . Les générateurs d'homologie réduite au niveau 0, c.-à-d. les composantes connexes (cc), sont diminuées de un. Noter que les calculs sont faits au niveau de l'homologie réduite $H^\#$.

Figure 2.21 - Exemple d'une λ -cellule élémentaire qui affecte l'homologie du niveau $\lambda-1$



L'ajout de e_p^1 à M^x crée un générateur au niveau l pour M^y .

Figure 2.22 - Exemple d'une λ -cellule élémentaire qui affecte l'homologie du niveau λ

2.11 Conclusion

Comme nous venons de le voir, la théorie de Morse permet d'établir le lien entre les points critiques d'une fonction définie sur une variété et la topologie de celle-ci. Ainsi, à chaque apparition d'un point critique d'indice λ dans la fonction f on associe une λ -cellule. La topologie de la variété étudiée et par la même son homologie, seront affectées par le type de cellule et la façon de l'attacher. C'est ces constats qui ont inspirés l'apparition du descripteur de Morse qui use de l'évolution des générateurs d'homologie lors des différents niveaux M^x pour décrire la forme de l'objet étudié.

Chapitre 3

Descripteur de forme basé sur la théorie de Morse

3.1 Introduction

Le lien établi par la théorie de Morse, entre la topologie d'une paire (M_y, M_x) et les points critiques situés entre ces deux niveaux, est à la base de l'idée qui a servi à concevoir le *descripteur de forme basé sur la théorie de Morse*. En effet, ce descripteur décrit l'information relative aux régions critiques d'une fonction de Morse définie sur une forme. Et ce, par le calcul des changements de topologie entre différents niveaux qui coïncident avec l'apparition ou la disparition de points critiques. Dans ce chapitre nous donnons une définition formelle de ce descripteur et nous en présentons les différentes propriétés. Nous explorons des moyens pour améliorer la distance de mesure et le choix des fonctions de Morse. Une dernière partie est consacrée à décrire une batterie de tests qu'on a appliqué.

3.2 Descripteur de Morse

Nous construisons une représentation topologique de la variété M à partir de la relation entre le changement de topologie de M et les points critiques d'une fonction de Morse φ définie sur M . La représentation sera donnée en terme de nombres Betti (les rangs de groupes d'homologie relative) d'une collection de paires de sous-ensembles de niveau (M_y, M_x) de φ obtenus pour toutes les valeurs de l'intervalle $I_\varphi = \left[\min_{q \in M} \varphi(q), \max_{q \in M} \varphi(q) \right]$ de valeurs de φ . D'une manière formelle notre descripteur appelé *Descripteur de Morse* est défini comme suit

Définition 3.1 Le *Descripteur de Morse* associé à la paire (M, φ) est la fonction

$MD_\varphi : \mathbb{R}^2 \times \mathbb{N} \rightarrow \mathbb{N}$ définie par :

$$MD_\varphi(x, y, k) = \begin{cases} \text{Rang}(H_k(M_y, M_x; \mathbb{R})), & \text{si } y \in I_\varphi \text{ et } x \leq y, \\ 0 & \text{Sinon.} \end{cases}$$

Vu que M est de dimension n , alors pour tout $x \leq y$, seuls les groupes d'homologie relative $H_k(M_y, M_x; \mathbb{R})$ ayant $0 \leq k \leq n$ seront possiblement non-triviaux. L'intervalle I_φ est fini vu que M est compact. Une discrétisation finie de l'intervalle I_φ (qui est fini puisque M est compact) produit suffisamment d'information topologique sur M pour pouvoir construire une représentation de sa forme topologique. Ce qui nous permet de dire que cet outil peut être utilisé dans la pratique comme montré dans la section suivante.

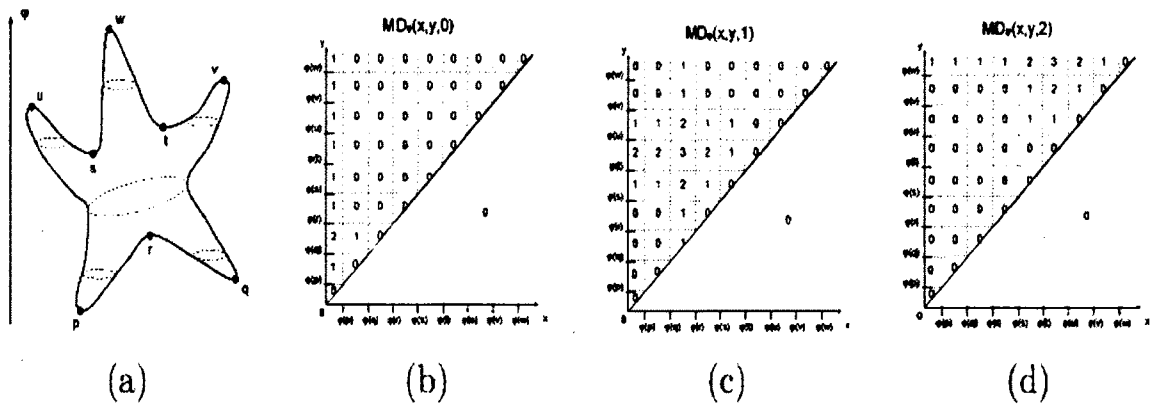


Figure 3.1 - Exemple d'une variété bidimensionnelle et la fonction hauteur discrétisée en neuf niveaux, ainsi que le descripteur de Morse associé défini par trois matrices

Exemple 3.1 La Figure 3.1 décrit une surface et les trois matrices représentant le descripteur de Morse $MD_\varphi(x, y, k)$, $k = 0, 1, 2$, tel que φ est la fonction hauteur sur la surface. Cette fonction hauteur est discrétisée de telle façon que le premier niveau est en dessous du minimum de la fonction. De cette manière le sous-groupe de niveau correspondant sera vide.

Les huit autres niveaux coïncideront avec l'apparition des huit points critiques de la fonction hauteur sur la surface. Dans la matrice (b), la première colonne donne le nombre de composantes dans les différents sous-groupes de niveau parcourus. Les entrées dans la seconde colonne correspondent aux nombres de composantes relatives dans les niveaux relatifs au niveau contenant le point critique p . En d'autres termes, les composantes différentes de celles générées par le premier point critique p . De même pour les tunnels et les cavités dans les matrices (c) et (d).

La Figure 3.2 montre un exemple concret des matrices associées à la forme et à la fonction de Morse utilisée

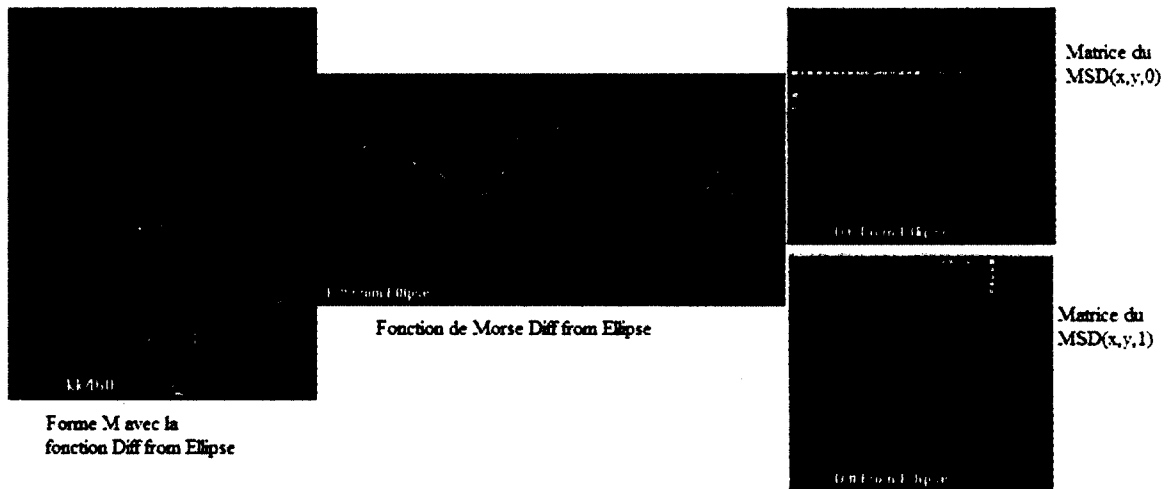


Figure 3.2 – Exemple d'une forme type contour et les matrices associées du descripteur de forme MSD

On a vu dans les théorèmes 2.5 et 2.6 (voir aussi [53]) que l'apparition d'un point critique p de φ avec λ comme indice, entre les niveaux M_x et M_y correspond topologiquement à l'attachement d'une λ -cellule e^λ par une fonction d'attachement f_p au niveau M_x . On y voit aussi une description homologique de cette transition topologique basée sur la fonction induite en homologie par la fonction d'attachement f_p . Cette information est, cependant,

intrinsèque à l'évolution globale des groupes d'homologie $H_k(M_x; \mathbb{R})$, $x \in I_\varphi$ tel que décrit dans [54]. Dans un même contexte, où un unique point critique, d'indice λ , est présent entre les niveaux M_x et M_y , le descripteur de Morse donnera une caractérisation plus affinée. En effet, par la propriété d'excision de l'homologie relative (voir [40], p. 82), on a

$$\begin{aligned} H_k(M_y, M_x; \mathbb{R}) &\cong H_k(M_x \cup_{f_p} e^\lambda, M_x; \mathbb{R}) \cong H_k(e^\lambda, \partial e^\lambda; \mathbb{R}) \\ &\cong \begin{cases} \mathbb{R}, & \text{si } k = \lambda \\ 0, & \text{sinon} \end{cases} \end{aligned}$$

Ainsi, $\text{MD}_\varphi(x, y, k) = 1$ si $k = \lambda$ et $\text{MD}_\varphi(x, y, k) = 0$ sinon. Dans un contexte plus général, le descripteur de Morse donnera directement une limite inférieure lors de la transition topologique entre n'importe quel paire de niveaux. En effet, en rajoutant un nouveau niveau M_z entre M_x et M_y , on peut engendrer plus d'information topologique. Plus précisément, on peut prouver le résultat suivant.

Théorème 3.1 Soit le triplet (M_y, M_z, M_x) de sous niveaux relatifs à une fonction de Morse $\varphi: M \rightarrow \mathbb{R}$ tels que $M_x \subseteq M_z \subseteq M_y$ alors

$$\text{MD}_\varphi(x, y, k) \leq \text{MD}_\varphi(x, z, k) + \text{MD}_\varphi(z, y, k) \quad \forall k \in \mathbb{N}$$

Preuve. Le triplet (M_y, M_z, M_x) produit la séquence exacte d'homologie suivante

$$\cdots \xrightarrow{\partial_*} H_k(M_z, M_x; \mathbb{R}) \xrightarrow{i_*} H_k(M_y, M_x; \mathbb{R}) \xrightarrow{j_*} H_k(M_y, M_z; \mathbb{R}) \xrightarrow{\partial_*} H_{k-1}(M_z, M_x; \mathbb{R}) \xrightarrow{i_*} \cdots$$

Et vu que nous utilisons l'homologie avec des coefficients réels, tous les groupes d'homologie et d'homologie relative dans cette séquence sont des espaces vectoriels. De par l'exactitude même de celle-ci, on peut écrire,

$$\begin{aligned}
H_k(M_y, M_x; \mathbb{R}) &\cong \ker j_{*k} \oplus \operatorname{im} j_{*k} \\
&\cong \operatorname{im} i_{*k} \oplus \ker \varphi_{*k} \quad (\text{par exactitude})
\end{aligned}$$

Etant donné que $\ker j_{*k}$ et $\operatorname{im} j_{*k}$ sont, respectivement, isomorphes à des sous-espaces des espaces vectoriels $H_k(M_z, M_x; \mathbb{R})$ et $H_k(M_y, M_z; \mathbb{R})$ il s'en suit que

$$\operatorname{MD}_\varphi(x, y, k) \leq \operatorname{MD}_\varphi(x, z, k) + \operatorname{MD}_\varphi(z, y, k)$$

3.3 Propriétés et calculabilité du descripteur de Morse

La motivation première du développement du descripteur de Morse MD est de l'utiliser comme un descripteur de forme. Nous devons donc spécifier sous quelles conditions le MD est invariant à des transformations rigides et à des changements d'échelles.

3.3.1 Invariance

Etant donné deux sous-variétés M et N de \mathbb{R}^n telles que $N = T(M)$, T étant une transformation rigide ou d'échelle (T , en fait, est assimilé à un difféomorphisme). Etant donné aussi $h: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction, on définira que h est invariante selon T si et seulement si $h_M = h_N \circ T$. On peut facilement établir que le MD possède les propriétés d'invariance de h . Plus précisément on donne la proposition suivante.

Proposition 3.1 Soient M et N deux sous-variétés fermées de \mathbb{R}^n telles que $N = T(M)$ pour un certain difféomorphisme T . Soit $h: \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction invariante sous T et telle que $f = h_M$ et que $g = h_N$ soient respectivement des fonctions de Morse sur M et N . Alors les descripteurs de Morse associés à la paire (M, f) et (N, g) coïncident.

Preuve. On rappelle que pour une paire (M, φ) consistant en une sous-variété fermée de \mathbb{R}^n et une fonction de Morse $\varphi: M \rightarrow \mathbb{R}$, le descripteur de Morse sera décrit tel que dans la définition 3.1 par.

$$\text{MD}_\varphi(x, y, k) = \begin{cases} \text{Rang}(H_k(M_y, M_x; \mathbb{R})), & \text{si } y \in I_\varphi \text{ et } x \leq y, \\ 0 & \text{Sinon.} \end{cases}$$

où, $I_\varphi = \left[\min_{q \in M} \varphi(q), \max_{q \in M} \varphi(q) \right]$. Observons que pour chaque $q \in N$, il existe un unique $p \in M$ tel que $q = T(p)$ et donc $f(p) = g \circ T(p) = g(q)$. Il s'en suit que $I_f = I_g$. En plus et par la propriété d'invariance, on déduit que pour tout $t \in \mathbb{R}$,

$$N_t^g = \{q \in N : g(q) \leq t\} = T(\{p \in M : f(p) \leq t\}) = T(M_t^f)$$

Ainsi, pour $x, y \in \mathbb{R}$ avec $y \in I_f$ et $x \leq y$, T va transformer de manière homéomorphe la paire (M_y^f, M_x^f) en la paire (M_y^g, M_x^g) . Ce qui induit l'isomorphisme dans l'homologie relative suivant

$$H_k(M_y, M_x; \mathbb{R}) \cong H_k(N_y, N_x; \mathbb{R}), \quad \forall k \in \mathbb{N}$$

Il s'ensuit que $\text{MD}_f(x, y, k) = \text{MD}_g(x, y, k)$ pour tout $(x, y, k) \in \mathbb{R}^2 \times \mathbb{N}$.

En utilisant, donc, des fonctions de Morse invariantes à des transformations rigides et à des changements d'échelles, le descripteur de Morse MD dérivé va hériter de ces propriétés d'invariances et il pourra, ainsi, être considéré comme un descripteur de forme. On le renommera donc *Descripteur de forme de Morse* noté *MSD* (pour *Morse Shape Descriptor*).

Remarque Plusieurs fonctions de Morse telles que la fonction hauteur et la distance à un point fixé (telle que la distance au centre de masse de la forme donnée) sont invariantes aux translations et aux rotations mais ne le sont pas aux changements d'échelle. Un moyen simple

de les rendre invariantes à cette transformation, tout en gardant l'invariance aux translations et aux rotations, serait de remplacer notre fonction φ par sa fonction normalisée correspondante

$$\bar{\varphi}(x) = \frac{\varphi(x) - \varphi_{\min}}{\varphi_{\max} - \varphi_{\min}}$$

ou φ_{\max} et φ_{\min} sont respectivement le maximum et le minimum des φ .

3.3.2 Relation avec des descripteurs topologiques connus

En autorisant le choix de valeurs de x telles que $x < \min_{q \in M} \varphi(q)$ et ainsi $M_x = \emptyset$, on pourra inclure dans notre description les paires (M_y, \emptyset) , ainsi, $H_k(M_y, M_x; \mathbb{R}) = H_k(M_y, \emptyset; \mathbb{R}) = H_k(M_y; \mathbb{R})$ pour tout $k \in \mathbb{N}$ et tout $y \in I_\varphi$. Le descripteur de Morse mesurera donc, lui aussi, l'évolution des groupes d'homologie de M tel que décrit dans [54], l'information ne sera, par contre, pas codée en une structure de graphe.

Un autre descripteur topologique de variétés unidimensionnelles connu, relié à notre descripteur, est le concept de fonction de taille introduit dans [55,56]. Etant donné une paire (M, φ) telle que décrite avant et y une valeur réelle, un relation d'équivalence \mathcal{R}_y est définie sur M tel que : deux points p et q seront \mathcal{R}_y -équivalents si $p=q$ ou qu'il existe un chemin continu entre p et q entièrement contenu dans M_y . Etant donné deux valeurs réelles x et y , la fonction de taille à valeurs entières $\ell_\varphi(x, y)$ du contour M consiste en le nombre de classes d'équivalence dans lesquelles l'ensemble M_x est divisé par la relation d'équivalence \mathcal{R}_y . En d'autres termes, $\ell_\varphi(x, y)$ est définie par $+\infty$ si $y < x$ et comme étant le nombre de composantes connexes de l'ensemble M_y qui contient au moins un point de M_x , sinon. Des algorithmes de composantes connexes peuvent être utilisés pour le calcul de la fonction de taille. À cette fin, lorsque $y < x$, le rang du groupe d'homologie relative $H_0(M_y, M_x; \mathbb{R})$ -qui est le nombre de composantes connexes de l'ensemble M_y qui ne contiennent aucun

point de M_x - ce rang donc, pourra être utilisé. La fonction de taille sera donnée par la différence entre le rang du groupe d'homologie d'ordre 0 soit $H_0(M_y; \mathbb{R})$ et le groupe d'homologie relative, d'ordre 0 lui aussi, $H_0(M_y, M_x; \mathbb{R})$ (voir [47]).

En normalisant, les valeurs de φ pour qu'elles appartiennent à l'intervalle $[0,1]$, de telle façon que $0 < \min_{x \in M} \varphi(x)$, on peut écrire comme avant $H_0(M_y; \mathbb{R}) = H_0(M_y, \emptyset; \mathbb{R}) = H_k(M_y, M_0; \mathbb{R})$. Ainsi, la fonction de taille sera obtenue à partir du descripteur de Morse comme suit

$$\ell_\varphi(x, y) = \text{MD}_\varphi(0, y, 0) - \text{MD}_\varphi(x, y, 0)$$

On peut voir notre descripteur tel un descripteur fonction de taille raffiné. Pour une variété à une dimension, notre descripteur donne exactement la même information que celle donnée par la fonction de taille. Malgré le fait que la théorie de la fonction de taille soit formulée, à l'origine, pour n'importe quelle variété, il n'y a pas de méthode formelle pour son calcul sur des formes de dimensions supérieures.

Remarque Dans nos suppositions, nous avons imposé des variétés fermées (i.e., connexes, compactes et sans frontières) pour nous garantir de travailler sur des intervalles fermés qui seront les domaines de nos fonctions de Morse, et ainsi, nous garantir que les générateurs d'homologie seront directement reliés aux points critiques de nos fonctions de Morse. Dans les cas où la variété possède une frontière, certains générateurs d'homologie seront reliés à celle-ci au lieu qu'à un point critique donné. En pratique, on pourra aussi gérer les surfaces connexes ouvertes et ce en exigeant de notre fonction de Morse d'avoir un domaine fini (délimité). La connexité garantit que les domaines soient des intervalles et l'ouverture garantie que les surfaces n'aient pas de frontières. Par exemple, pour des surfaces qui sont de

graphes de fonctions lisses à deux variables définies sur de domaines connexes et délimités, la seule frontière en question sera la frontière extérieure qui pourra être ignorée (où réduite à un point) pour obtenir une surface fermée.

3.3.3 Calculabilité du MSD

Un problème se pose lors d'expérimentations dans lesquelles nous calculons le MSD en utilisant des fonctions (qu'on appellera fonctions de mesure) qui ne sont pas nécessairement des fonctions de Morse. Notre première tentative d'étendre la définition du MSD à des fonctions continues a été d'utiliser la même approche que dans ([55], proposition 1.1).

Une observation importante ici est que la combinaison de résultats classiques dans des analyses réelles et les résultats dans [51] (voir Corollaire 6.8) et [57] (voir Théorème 2.20), cette combinaison, disons-nous, montre que n'importe quelle fonction réelle et continue sur une variété fermée et lisse peut être approximée uniformément par une fonction de Morse. Il est démontré dans [57] que ceci peut être fait en appliquant des perturbations linéaires sur chaque système de coordonnées local de la variété.

L'approche dans [55] peut être transcrite à notre problème en prouvant qu'étant donné deux fonctions de Morse $f, g : M \rightarrow \mathbb{R}$ telles que $\|f - g\|_\infty < \varepsilon$, ou

$$\|f - g\|_\infty = \max_{m \in M} |f(m) - g(m)|,$$

les descripteurs de Morse associés aux paires (M, f) et (M, g) seront proches d'une certaine manière. On pourrait, par exemple, essayer de prouver, comme dans [55], que pour tout $x, y \in \mathbb{R}$ avec $x < y$

$$\text{MSD}_f(x - \varepsilon, y + \varepsilon, k) \leq \text{MSD}_g(x, y, k) \leq \text{MSD}_f(x + \varepsilon, y - \varepsilon, k),$$

qui est la formule établie dans [55] pour les fonctions de taille. Pour le MSD pourtant, ces inégalités ne sont pas généralement vraies, à moins d'imposer certaines contraintes sur f et g . Ce fait est illustré dans la figure suivante.

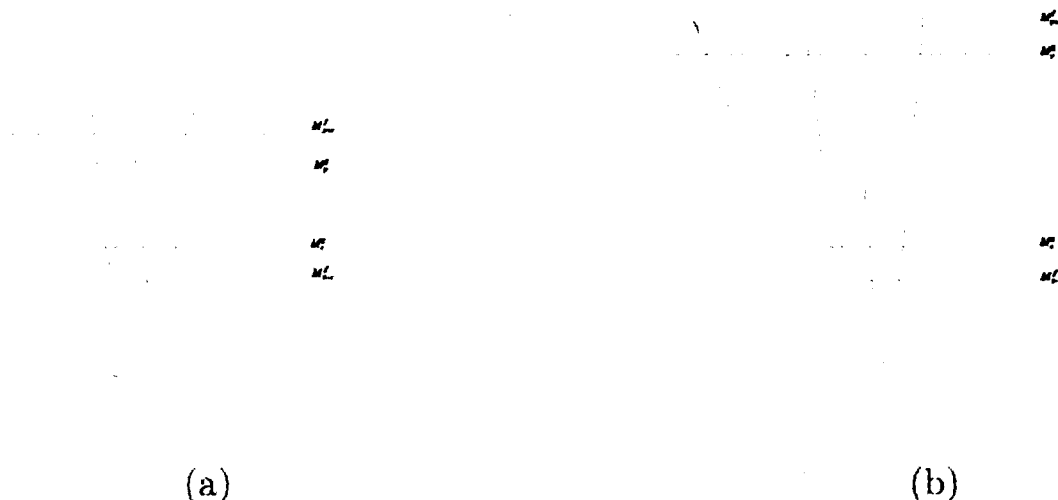


Figure 3.3 - dans (a) il est facile de voir que $\text{MSD}_f(x-\varepsilon, y+\varepsilon, 0) = 1$ alors que $\text{MSD}_g(x, y, 0) = 0$. Dans (b) $\text{MSD}_f(x-\varepsilon, y+\varepsilon, 0) = 0$ alors que $\text{MSD}_g(x, y, 0) = 1$. Mêmes observations à propos de $\text{MSD}_f(x+\varepsilon, y-\varepsilon, 0)$ et $\text{MSD}_g(x, y, 0)$.

Observons, tout d'abord, que le MSD est défini pour toute fonction de mesure continue. L'unique intérêt dans l'utilisation de fonctions de mesure de Morse réside dans le fait que l'on peut construire une connexion entre le MSD et les points critiques de la fonction de mesure de Morse. Tel que mentionné précédemment, étant donné une fonction $g : M \rightarrow \mathbb{R}$ continue et définie sur une variété lisse, pour tout $\varepsilon > 0$, il existe une fonction de Morse $f : M \rightarrow \mathbb{R}$ telle que $\|f - g\|_\infty < \varepsilon$. Nous pouvons choisir un ε très petit de sorte que les topologies des niveaux de f et g , dont la différence est de ε , soient similaires. On aura donc le résultat suivant

Proposition 3.2 Supposons que M est une variété lisse et fermée, $g : M \rightarrow \mathbb{R}$ une fonction de mesure continue et $f : M \rightarrow \mathbb{R}$ une fonction de Morse telle que $\|f - g\|_\infty < \varepsilon$ et pour tout $x, y \in \mathbb{R}$ avec $x + \varepsilon < y - \varepsilon$ les niveaux M_x et M_y seront des déformations-rétraction de $M_{x+\varepsilon}$ et $M_{y+\varepsilon}$ respectivement, ainsi que les niveaux $M_{x-\varepsilon}$ et $M_{y-\varepsilon}$ seront, eux aussi, des déformations-rétraction des niveaux M_x et M_y , respectivement. On aura alors

$$\text{MSD}_f(x - \varepsilon, y + \varepsilon, k) = \text{MSD}_g(x, y, k) = \text{MSD}_f(x + \varepsilon, y - \varepsilon, k).$$

Preuve. Soit $\varepsilon > 0$ tel que $x + \varepsilon < y - \varepsilon$. Il s'en suit que

$$M_{x-\varepsilon}^f \subset M_x^g \subset M_{x+\varepsilon}^f \subset M_{y-\varepsilon}^f \subset M_y^g \subset M_{y+\varepsilon}^f$$

A partir du triplet $M_{y+\varepsilon}^f \supset M_y^g \supset M_x^g$ on pourra produire la séquence exacte d'homologie suivante

$$\cdots H_{k+1}(M_{y+\varepsilon}^f, M_y^g) \xrightarrow{\partial_*} H_k(M_y^g, M_x^g) \xrightarrow{i_*} H_k(M_{y+\varepsilon}^f, M_x^g) \xrightarrow{j_*} H_k(M_{y+\varepsilon}^f, M_y^g) \xrightarrow{\partial_*} H_{k-1}(M_y^g, M_x^g) \xrightarrow{i_*} \cdots$$

Vu que M_y^g est une déformation-rétraction de $M_{y+\varepsilon}^f$, il s'en suit par la proposition 2.1 que $H_{k+1}(M_{y+\varepsilon}^f, M_y^g)$ et $H_k(M_{y+\varepsilon}^f, M_y^g)$ sont des groupes triviaux. Par exactitude, on a

$$\ker i_{*k} = \text{im } \partial_{*k} = 0 \text{ et que } \text{im } i_{*k} = \ker j_{*k} = H_k(M_{y+\varepsilon}^f, M_x^g)$$

Il s'en suit que i_{*k} est un isomorphisme pour tout $k \in \mathbb{N}$ et ainsi

$$H_k(M_{y+\varepsilon}^f, M_x^g) \cong H_k(M_y^g, M_x^g)$$

Les mêmes arguments appliqués au triplet $M_{y+\varepsilon}^f \supset M_x^g \supset M_{x-\varepsilon}^g$ donneront la relation

$$H_k(M_{y+\varepsilon}^f, M_{x-\varepsilon}^g) \cong H_k(M_{y+\varepsilon}^f, M_x^g),$$

Et finalement le fait que

$$H_k(M_{y+\varepsilon}^f, M_{x-\varepsilon}^f) \cong H_k(M_y^g, M_x^g)$$

démontre la première égalité de la proposition. En utilisant des arguments similaires, on pourra prouver que

$$H_k(M_{y-\varepsilon}^f, M_{x+\varepsilon}^f) \cong H_k(M_y^g, M_x^g)$$

ce qui démontre la seconde égalité de la proposition.

Notons que certaines simplifications peuvent être déduites sur une variété, notamment selon la proposition suivante

Proposition 3.3 Soit (M_y, M_z, M_x) un triplet de complexes tels que $M_x \subseteq M_z \subseteq M_y$ et $\varphi: M \rightarrow \mathbb{R}$ une fonction de Morse. Si $H_k(M_y M_z) \cong 0, \forall k \in \mathbb{N}$, alors $MSD_\varphi(x, y, k) = MSD_\varphi(x, z, k)$. Pareillement, si $H_k(M_z M_x) \cong 0, \forall k \in \mathbb{N}$, alors $MSD_\varphi(x, y, k) = MSD_\varphi(z, y, k)$.

3.4 Algorithme

L'algorithme utilisé pour calculer le *descripteur de Morse* donne en terme de temps de calcul un ordre de $\Theta(N^2)$ calculs d'homologie relative où N spécifie la résolution des matrices du MSD_φ .

Entrées: \mathcal{M} : complexe cubique
 $\hat{\phi}$: fonction de Morse discrete sur \mathcal{M}
 N : resolution de \widehat{MSD}_φ

Sorties: \widehat{MSD}_φ : descripteur de Morse pour la forme discretisé

pour $\lambda \in \{0, \dots, \dim(\mathcal{M})\}$ **faire**
 $\widehat{MSD}_\varphi(\cdot, \cdot, \lambda) \leftarrow 0;$
 $d\hat{\phi} \leftarrow (\hat{\phi}^{\max} - \hat{\phi}^{\min})/N;$
pour $(y, \hat{y}) \leftarrow (\hat{\phi}^{\min}, 0)$ **jusqu'à** $(\hat{\phi}^{\max}, N-1)$ **par** **increment** de $(d\hat{\phi}, 1)$
faire
pour $(x, \hat{x}) \leftarrow (\hat{\phi}^{\min}, 0)$ **jusqu'à** (y, \hat{y}) **par** **increment** de $(d\hat{\phi}, 1)$ **faire**
 $\widehat{MSD}_\varphi(\hat{x}, \hat{y}, \lambda) \leftarrow \text{Rang } H_\lambda(M^y, M^x);$

Algorithme de calcul du descripteur de Morse pour la forme

Notons qu'en modifiant l'ordre de parcours des paires (M^y, M^x) , on pourra pleinement profiter des simplifications de calcul que nous offre la proposition 3.3 qui se produisent lorsqu'une des paires (M^y, M^x) ou (M^x, M^y) est triviale.

3.5 Mesure de distance

Soient M et N les deux complexes cubiques à comparer. Et soient MSD_φ^M et MSD_φ^N leurs descripteurs de Morse respectifs. En considérant que la fonction de Morse φ est normalisée

pour produire ses valeurs dans l'intervalle ouvert $(0,1)$, une distance \mathcal{D}_{MSD} peut être alors définie comme suit

$$\mathcal{D}_{MSD}(MSD_{\varphi}^M, MSD_{\varphi}^N) = \sum_{\lambda=0}^{\dim} \left(\sum_{j=0}^{n-1} \sum_{i=0}^{n-1} |MSD_{\varphi}^M(i, j, \lambda) - MSD_{\varphi}^N(i, j, \lambda)| \right)$$

Outre le problème de la persistance des points critiques qui sera adressé dans un prochain paragraphe, cette distance nous pose le problème de sa sensibilité au bruit, dont une des principales conséquences est l'apparition de valeurs aberrantes ou extrêmes dans certaines cases des matrices. Ce problème fausse la comparaison, ce qui nous a poussé à vouloir éliminer le poids des ces valeurs aberrantes en pondérant chaque case utilisée dans le calcul des différences, par un facteur qui sanctionnerait les valeurs aberrantes et ce, selon la méthode de calcul suivante

- Calculer la valeur médiane de chaque vecteur des matrices.
- Pondérer chaque élément de matrice x_i dans le calcul de la différence par

$$\frac{1}{|x_i - Mediane|}.$$

Illustrons cette méthode en l'appliquant sur un exemple

Exemple pour les deux matrices de dimensions 3 $A = \begin{pmatrix} x_1 & x_4 & x_7 \\ x_2 & x_5 & x_8 \\ x_3 & x_6 & x_9 \end{pmatrix}$ et $B = \begin{pmatrix} y_1 & y_4 & y_7 \\ y_2 & y_5 & y_8 \\ y_3 & y_6 & y_9 \end{pmatrix}$

on calcule les six valeurs médianes des vecteurs colonne $Me_1^1 = Mediane(x_1, x_2, x_3)$,
 $Me_2^1 = Mediane(y_1, y_2, y_3)$, $Me_1^2 = Mediane(x_4, x_5, x_6)$, $Me_2^2 = Mediane(y_4, y_5, y_6)$,
 $Me_1^3 = Mediane(x_7, x_8, x_9)$ et $Me_2^3 = Mediane(y_7, y_8, y_9)$.

Les différences d_i avec $i = 1; 9$ seront alors calculées comme suit

$$d_i = \left| \left(\left(\frac{1}{|x_i - Me_1^k|} \right) * x_i \right) - \left(\left(\frac{1}{|y_i - Me_2^k|} \right) * y_i \right) \right|$$

avec $k=1$ pour $i=1;3$, $k=2$ pour $i=4;6$ et $k=3$ pour $i=6;9$. On sanctionnera de cette façon les valeurs extrêmes en réduisant leurs poids dans le calcul de la différence.

Problème de la persistance

On peut intuitivement deviner que les points critiques de φ n'ont pas tous la même contribution dans le calcul des distances. La contribution de chacun, est proportionnelle à son importance relative par rapport à celles des autres points critiques. L'importance relative d'un point critique peut être définie comme, le degré du lissage de la forme requis, pour qu'il disparaisse. La Figure 3.4 [45] illustre deux exemples de cette notion de persistance.

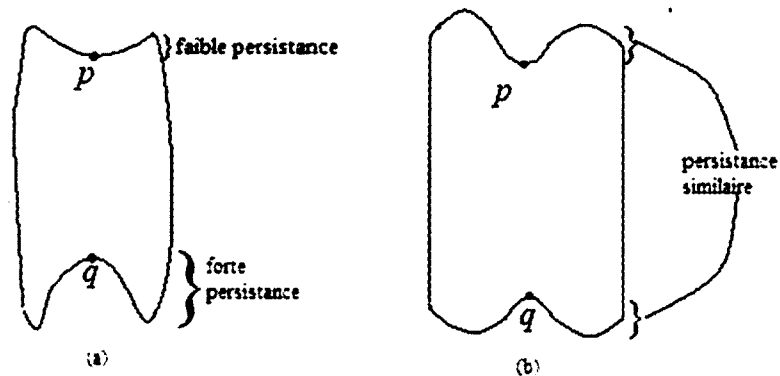


Figure 3.4 – Exemple de la notion de persistance

Bien qu'elle donne des résultats très concluants, notre distance ne traite pas ce problème et donne ainsi la même contribution aux points p et q de la Figure 3.5, alors que leurs importances relatives (persistances) sont différentes. Mais le traitement de la persistance est davantage une volonté d'optimisation qu'un réel problème structurel dans notre cas

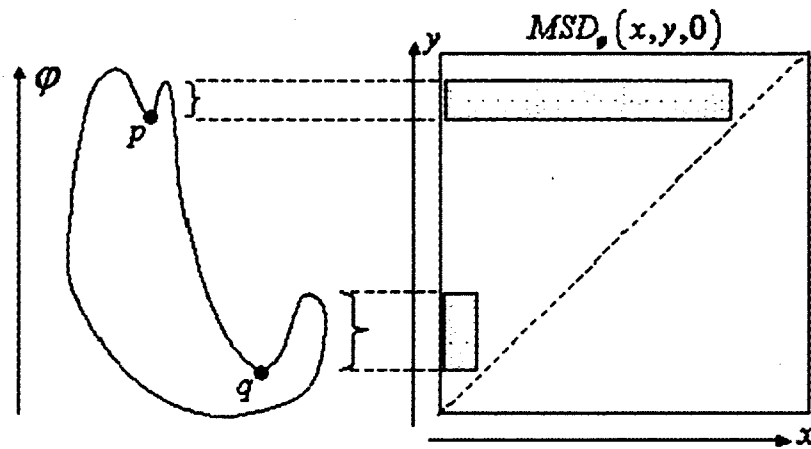


Figure 3.5 – Exemple de la persistance versus contribution.

S'inspirant de la définition de l'importance relative d'un point critique, nous avons pensé à associer à celui-ci une pondération normalisée, équivalente au nombre de lissages nécessaires pour faire disparaître ce point critique. Ce traitement, dépendant de la résolution des matrices, serait effectué comme prétraitement sur chaque image de la base de données. Mais il s'avère qu'étant donné le nombre de points critiques à traiter, même à de faibles résolutions, le coût en terme de temps de calcul devient trop élevé pour penser à implanter une telle méthode.

Une autre méthode inspirée de l'*Analyse en Composantes Principales* (ACP) a été testée dans un but de réduction de dimensionnalité et de sélection des variables les plus pertinentes à comparer. Ainsi, pour chaque matrice à notre disposition, on a considéré les lignes comme les occurrences et les colonnes comme les variables. On applique tout le processus de l'ACP : Normalisations et réduction des matrices, calcul des matrices de variances-covariances, calcul des vecteurs et valeurs propres, projections sur les nouvelles variables qui par définition sont de moindre dimension (ces calculs ont été faits sur un logiciel spécialisé : SPSS en l'occurrence). Outre la lourdeur de cette méthode, le fait que l'ACP soit appliquée à toutes les matrices d'une même forme (trois dans notre cas) supprimera tout lien logique entre celles-ci (lien entre la dimension 0, 1 et 2). Le fait aussi que toutes les lignes de nos matrices ne correspondent pas forcément avec l'apparition d'un point critique mais des fois correspondent avec l'apparition de plusieurs sur une même ligne, ce fait, disons-nous, rend

très floue la signification d'une ligne (utilisée comme occurrence dans l'ACP) et très floue la signification des nouvelles variables créées après l'ACP. Pour finir de nous décourager d'implanter cette méthode, il est arrivé qu'après avoir appliqué l'ACP aux deux matrices à comparer, on obtienne des dimensions différentes pour les deux résultats.

Afin de mieux cerner cette question de la persistance des points critiques, on pourrait suggérer, pour des travaux futurs, l'étude de la relation entre les points critiques d'une même dimension mais aussi de dimensions différentes. Mais aussi explorer la persistance homologique [58].

3.6 Combinaison des fonctions de Morse

Commençons tout d'abord par un constat, celui que le pouvoir de caractérisation de l'outil descripteur de Morse est largement dépendant d'un choix adéquat des fonctions de Morse φ . Comme expliqué dans le premier chapitre, pour bien caractériser les formes, le descripteur de Morse devrait demeurer similaire pour des formes similaires et être assez différent pour des formes différentes. Dans la Figure 3.6 on illustre ce problème de mauvais choix de la fonction de mesure (hauteur par rapport à l'axe mineur), où on obtient quasiment le même descripteur pour des formes assez différentes.

Malheureusement, la théorie que nous avons développée jusqu'à présent pour le descripteur de forme de Morse ne donne pas d'outils pour déterminer la manière avec laquelle on choisit des fonctions de Morse adéquates pour la description d'une classe de formes donnée. Le MSD est très dépendant de cette fonction de Morse utilisée lors de sa définition et il n'existe pas de système fourni assez développé pour comparer des représentations du MSD d'une même forme mais basées sur deux fonctions de Morse différentes. Ainsi à ce stade, le choix d'une fonction de Morse adéquate dépend de l'application et se fait de manière heuristique. Nous présenterons dans la suite la méthode heuristique qui nous permet, entre autre, de choisir la fonction de Morse (parmi plusieurs) la plus adéquate pour une classe de formes.

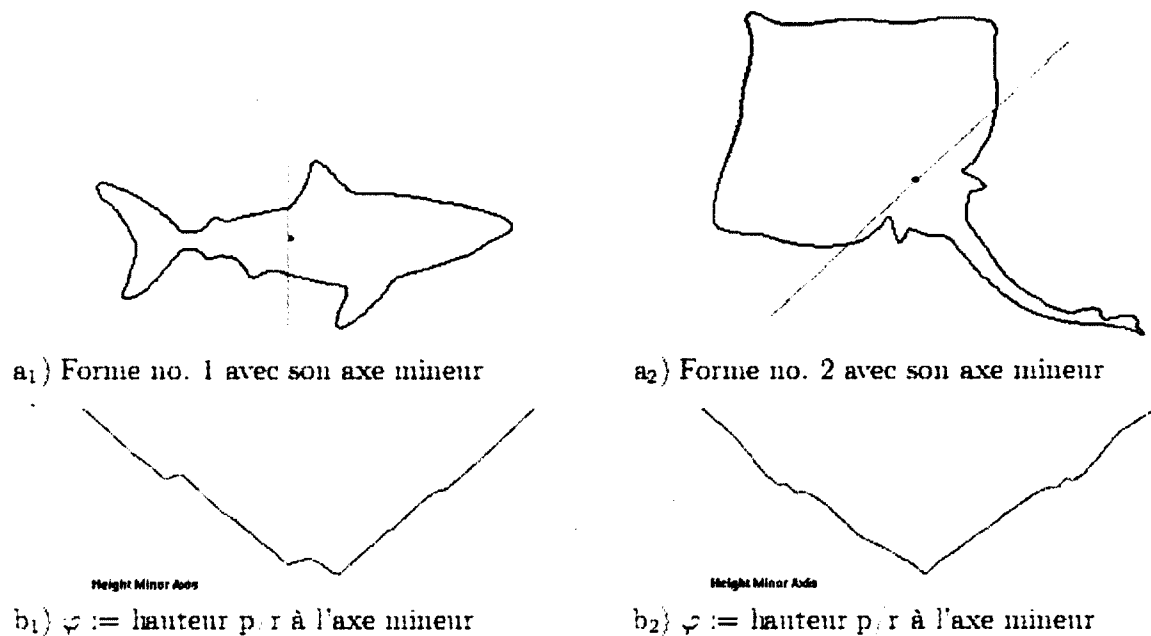


Figure 3.6 - Importance du choix adéquat de la fonction de mesure

Quatre fonctions de Morse ont été utilisées : la distance par rapport au centre de masse, la distance par rapport à l'axe majeur, la différence par rapport à l'ellipse qui approxime le mieux la forme et le cosinus de l'angle formé entre le vecteur tangent à la forme et l'axe majeur. L'utilisateur peut choisir d'en utiliser une seule comme de les combiner pour une description plus précise. Cependant et comme expliqué précédemment, une fonction de Morse peut être pertinente pour un type de formes et ne pas l'être pour un autre type. Et lorsqu'on combine plusieurs fonctions pour calculer la distance entre deux formes, la simple somme des distances de chacune des fonctions peut fausser la comparaison car elle donne le même poids aux distances des différentes fonctions de Morse. C'est ce qui nous a poussé à chercher un moyen de pondérer ces distances en fonction de la pertinence de la fonction de morse.

Pour comprendre comment nous calculons les poids à donner aux distances lors d'une combinaison de plusieurs fonctions de Morse, il faut savoir ce qui, pour nous, constitue une fonction de Morse non pertinente pour une classe de formes donnée. En effet, il est évident

que lorsqu'une fonction donne deux descripteurs très proches pour deux formes très différentes, cette fonction n'est pas adaptée à cette comparaison. Mais si pour une même classe de formes (par définition qui se ressemblent) la fonction de Morse choisie donne des descripteurs trop similaires, ce qui implique que les distances entre ses formes, en utilisant cette fonction, seront très proches l'une de l'autre. Alors, on peut dire que cette fonction n'est pas pertinente pour des comparaisons intra-classe, car elle a une mauvaise séparabilité. En d'autres termes, on a remarqué que les distances retournées vont se concentrer autour de leur moyenne. À l'inverse, une fonction qui donne une distribution des distances moins concentrée autour de la moyenne serait plus pertinente pour comparer des formes ressemblantes car elle mesure plus de caractéristiques pour ce type de formes et devrait avoir un plus grand poids lors de la combinaison.

En pratique, nous appliquons à notre descripteur une phase d'apprentissage lors de laquelle les poids des distances à combiner seront précalculés pour les comparaisons futures. Les étapes de cet apprentissage sont les suivantes

- i. L'utilisateur choisit le nombre d'images à prendre de chaque classe pour effectuer l'apprentissage.
- ii. On choisit aléatoirement une image de chaque classe et on calcule ses distances avec les images choisies dans la première étape. Et ce, pour les quatre fonctions de Morse. Cette image sera considérée par la suite comme le représentant de la classe.
- iii. On génère la moyenne et l'écart type des distances pour chaque fonction.
- iv. On calcule les différentes distances entre les représentants de classes.

Par la suite, et étant donné une image requête, on désire la comparer avec notre base de données. La politique de calcul des pondérations va dépendre du type de comparaison que l'on a à faire : intra classe ou inter classes.

- Lors des comparaisons intra-classe, les fonctions de Morse sont classées selon leur degré de dispersion autour de la moyenne. Plus grand est l'écart type, meilleure est la pertinence de la fonction pour cette classe. On utilisera donc les écarts type, calculés à l'étape ii, comme les poids de chacune des distances à combiner. Il est évident que les poids sont normalisés de sorte que leur somme soit égale à 1.
- Pour les comparaisons inter-classe, étant donné les distances entre les deux représentants des deux classes concernées. Soient R_i et R_j les représentants des deux classes C_i et C_j . Supposons que l'on ait quatre fonctions de Morse f_1, f_2, f_3, f_4 à combiner. Soient les deux vecteurs $V_i(m_1^i, m_2^i, m_3^i, m_4^i)$ et $V_j(m_1^j, m_2^j, m_3^j, m_4^j)$ associés à R_i et R_j ; avec m_k^i la distance moyenne selon la fonction f_k , entre le représentant R_i et les images sélectionnées de sa propre classe (phase ii de l'apprentissage). Et soit enfin le vecteur $D_{ij}(d_1, d_2, d_3, d_4)$ des distances entre R_i et R_j . Lorsque pendant un traitement d'une requête on aura à comparer l'image requête qui appartient à la classe C_i avec une image de la classe C_j (et vice versa), le poids p_k de la fonction de Morse f_k , lors de la combinaison des distances sera calculé de la sorte

$$p_k = \frac{((d_k - m_k^i) + (d_k - m_k^j))}{2}$$
 avec $k = 1, 2, 3, 4$; On désire favoriser les fonctions qui donnent une distance inter-représentant supérieure à la moyenne des différences de chaque représentant avec les images de sa classes.






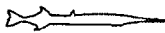






3.7 Présentation des résultats expérimentaux

Comme il est coutume de faire pour beaucoup de descripteurs de forme connus, un outil de CBIR (*Content-based image retrieval pour recherche d'images par le contenu*) a été développé pour valider en pratique le descripteur de forme de Morse. Notre outil suit le paradigme du *query-by-example* (QBE). Ce paradigme, dans le contexte des CBIR, consiste à fournir une image comme requête, extraire les propriétés pertinentes pour notre algorithme, mesurer la distance entre la requête et les images d'une base de données et finalement trier les images selon l'ordre croissant de leur distance par rapport à la requête. Le model QBE est très largement utilisé dans les systèmes de CBIR.

Nous avons utilisé la base de données SQUID fournie par [59]. Cette base de données est utilisée par le standard MPEG-7 pour l'évaluation des descripteurs de forme. Elle est aussi répertoriée dans la section Demo du site *Computer Vision Home Page* depuis 1997. Cette base consiste en 1100 images de poissons (sous forme de fichiers texte). La base de données a été manuellement classifiée en 11 classes. Pour chaque ensemble de quinze images ou plus qui se ressemblent, une classe a été créée. Les dix premières classes sont formées des images visuellement ressemblantes, la dernière classe contient toutes les images qui n'ont pas pu être classées, car ne ressemblant à aucune des dix autres classes. Le Tableau 3.1 montre un représentant de chacune des dix classes et deux images des la dernière classe

Le même groupe de recherche qui a construit la base de donnée SQUID a développé le descripteur de forme *curvature scale-space* (CSS) [60] qui a été adopté comme standard de représentation des formes dans MPEG-7. Le CSS utilise une paramétrisation de la courbure d'un contour qui est lissé par plusieurs filtres Gaussien de σ différents. L'image CSS d'un contour sera alors définie par le groupe de points (l, σ) pour lesquels la courbure (lissée) est à zéro et le descripteur CSS consistera en les maxima (pics) de l'image CSS. Ce même groupe fourni aussi un outil en ligne [61] qui permet d'appliquer le CSS sur la base de données SQUID, ce sera, entre autres, avec cet outil que nous comparerons notre descripteur.

Tableau 3.1 – Échantillons de chaque classe de notre base de données

Classe	Nb	Exemple	Classe	Nb	Exemple	Classe	Nb	Exemple
A	206		E	69		I	37	
B	144		F	59		J	17	
C	125		G	45		misc	265	
D	74		H	59				

Quatre types de tests ont été effectués sur notre descripteur pour en valider la puissance. Les mesures classiques de *précision* et *rappel* ont, bien sur, été utilisées avec les ensembles suivants, $A = \text{résultats pertinents}$; $B = \text{résultats non-pertinents}$ et $C = \text{images pertinentes exclues des résultats}$. Avec $\text{précision} = \frac{|A|}{|A| + |B|}$ et $\text{rappel} = \frac{|A|}{|A| + |C|}$.

Mais, à ce stade se pose la question de la définition de la pertinence ou pas d'un résultat. Ainsi, il est admis dans les expériences du même type qu'un résultat est pertinent si l'image appartient à la même classe que la requête. Mais nous présenterons dans la troisième expérience la notion de pertinence perceptuelle d'un résultat.

3.7.1 Première expérience

Nous nous inspirons dans cette expérience du standard **MPEG-7** en terme de méthode de test, et le *Core experiment CE-Shape-1* partie B, en l'occurrence [62], où on désire mesurer la performance de la recherche basée sur la similarité. Dans cette expérience, on teste notre descripteur avec un nombre réduit à 50 requêtes appartenant à différentes classes. Le nombre de requêtes par classe est proportionnel à l'importance de celle-ci en terme de nombre d'images. Mentionnons que le système donnera pour chaque requête 50 images en résultats ordonnées en ordre croissant de la distance avec la requête. On calcule la précision globale en faisant la moyenne de celles de toutes les requêtes. Figure 3.7 donne un aperçu de quelques résultats après deux requêtes.



Figure 3.7 - Aperçu de 5 résultats à deux requêtes dans notre outil

Après calcul, on obtient une précision moyenne de 73% et un rappel d'environ 26%. Cette précision est en fait inférieure à la précision réelle de notre outil. Ceci s'explique très facilement, par la méthode de calcul et à la disparité des cardinalités de nos classes. Ainsi si on remarque la classe J, elle compte 17 images, mais vu qu'on génère à chaque requête, 50 images résultats, même si notre outil sort les 16 autres images de la classe J à une requête de

cette même classe (ce qui donne 100% de bons résultats), le fait que l'on divise 17 par 50 (le nombre total de résultats), ce fait disons-nous, diminue à 34% la précision. Et lorsqu'on fait la moyenne des précisions pour toutes les classes, il est normal que cela baisse la moyenne globale des précisions. Si on prend la classe A par exemple, sa précision est de 78% (sa cardinalité élevée aidant à donner plus de résultats pertinents parmi les 50). La faiblesse du rappel est due à la raison inverse et ce car pour les classes à grande cardinalité (A, B par exemple) on ne prend que 50 résultats, ce qui fait augmenter le nombre C. Alors que si on calculait le rappel pour la classe J nous aurions un rappel beaucoup plus élevé (58%). Nous avons donc pensé à proposer d'autres tests sur notre descripteur.

3.7.2 Deuxième expérience

On compare maintenant les résultats de notre descripteur avec ceux donnés à travers l'outil standard de MPEG-7, qui applique l'algorithme CSS sur la base de données SQUID. Un résultat sera considéré pertinent pour une requête dans notre descripteur, si ce même résultat apparaît comme réponse à la même requête sur l'outil CSS. On applique donc 50 images requêtes aux deux descripteurs et on voit dans quel proportion notre outil donne les mêmes résultats que le standard MPEG-7. Notons que cet outil est conçu pour sortir un maximum de 18 résultats pour chaque requête. Les résultats ont été très concluants, car en moyenne dans les 18 résultats pour chacune des 50 requêtes sur CSS, on avait 76.5% de résultats semblables pour cette même requête sur notre outil. Il faut dire par contre que l'ordre dans lequel sont trouvés les résultats n'est pas le même dans les deux descripteurs.

3.7.3 Troisième expérience

Nous avons pensé ensuite à tester la pertinence perceptuelle des résultats fournis par notre outil. Pour cela nous avons fourni à trois étudiants un même ensemble de cinq images de notre base de données et nous leur avons demandé de donner pour chacune d'elles, neuf autres images qui lui ressemblaient le plus à partir de la base de données. Nous avons ensuite compilé les résultats pour ne garder que dix images par requête (selon notre jugement). Nous obtenons ainsi une sorte de nouvelle classification de notre base de données en 5 classes de 10 images chacune. Nous avons procédé par la suite à l'expérience suivante : on fournit à

notre système 5 requêtes parmi les 50 images sélectionnées plus haut. On prend les dix premiers résultats et on calcule la précision et le rappel. Ceux-ci sont ensuite schématisés dans la courbe de la Figure 3.8. Chaque point de la courbe est la moyenne des 5 requêtes. Le nombre total de points est 10 (i.e., on calcule la précision et le rappel pour les réponses contenant de 1 à 10 formes). Ainsi, le point en haut à gauche du diagramme correspond à la valeur du précision/rappel pour la meilleure réponse (*best match*), tandis que le point en bas à droite correspond au précision/rappel pour la réponse entière (les dix réponses rendues). Un résultat est pertinent s'il appartient à la même classe que la requête (créée lors du début de cette expérience).

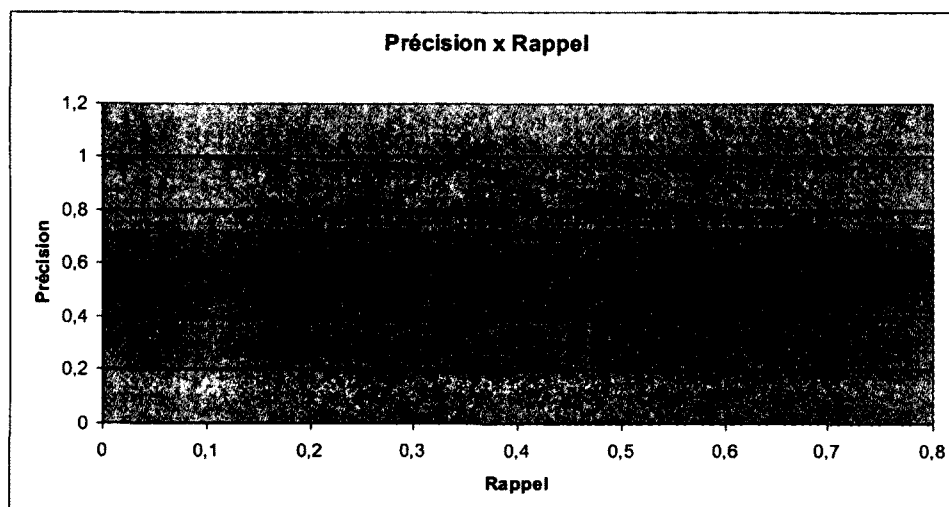


Figure 3.8 - Précision x Rappel : Plus haute est la courbe meilleur est le descripteur

Pour donner un peu un ordre d'idée entre notre descripteur dans cette expérience et d'autres descripteurs nous avons réimprimé de [63] la Figure 3.9 qui montre la performance des descripteurs de Fourier, les moments géométriques et les moments séquentiels (décrits dans l'article) pour la même expérience que la notre. On notera toutefois, que dans le cas de ces descripteurs le système reçoit 20 requêtes et sort 50 résultats. La pertinence y est aussi jugée par la perception humaine.

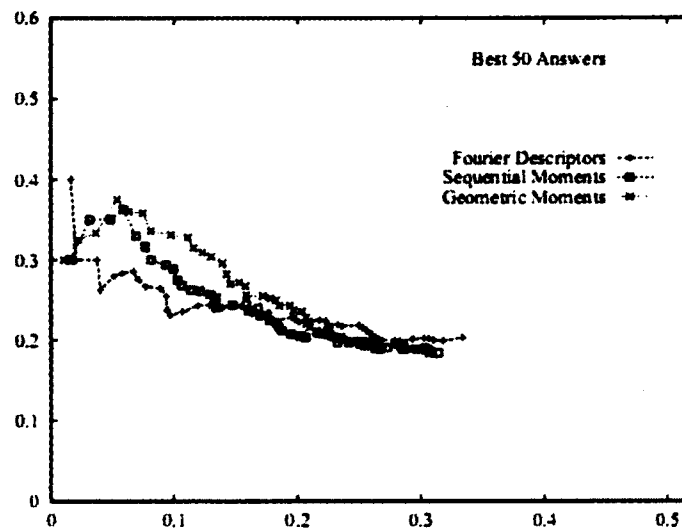


Figure 3.9 - Précision x rappel pour les descripteurs décrit ci-dessus

Il est évident que notre descripteur est meilleur sur tous les plans que ceux décrit dans la figure 3.9

3.7.4 Quatrième expérience

Nous voulons tester maintenant, l'apport de la phase d'apprentissage et la combinaison des fonctions de Morse tel que décrit précédemment en terme de précision de recherche. Pour les conditions du test, nous fournissons 5 requêtes de chaque classe à deux versions du descripteur de Morse. Dans la première les quatre distances (normalisées) des quatre fonctions de Morse sont sommées pour produire la distance entre deux formes. Tandis que dans la deuxième version, nous avons préalablement appliqué un apprentissage à notre descripteur (avec 20 images par classe). Les distances sont ensuite calculées en combinant celle de chaque fonction, selon l'algorithme décrit précédemment. Nous comptons le nombre d'images pertinentes dans les 50 résultats pour les deux cas. Comme dans la première expérience, un résultat est pertinent s'il appartient à la même classe que la requête. La précision moyenne sera ensuite calculée pour chaque classe selon qu'on ait utilisé l'apprentissage et la combinaison ou pas. Celles-ci sont arrondies et résumées dans le Tableau 3.2

Tableau 3.2 – Précisions moyennes avec et sans apprentissage

	A	B	C	D	E	F	G	H	I	J
Sans apprentissage	67%	86%	65%	92%	44%	73%	41%	89%	33%	26%
Avec apprentissage	71%	86%	68%	92%	68%	71%	67%	94%	49%	28%

Nous remarquons dans ce tableau une disparité dans l'évolution de la précision avant et après la combinaison. Les cas des classes A et C sont similaires car pour les deux, la précision est plus faible que dans la classe B qui est de dimension semblable. Cette faiblesse relative s'explique par le fait que les formes dans la classe A et C sont assez ressemblantes ce qui donne souvent des images de l'une mélangées avec les résultats d'une requête sur l'autre, d'où la baisse relative de précision. L'augmentation en précision avant et après la combinaison sur ces deux classes n'est pas négligeable. Les classes B, D et H ont de très bonnes précisions, ceci est dû au fait que ces classes sont très typées, dans le sens où leurs images respectives se ressemblent fortement entre elles et diffèrent significativement des images des autres classes. L'évolution de la précision dans les classes B, D, H F et J est insignifiante ou quasi nulle. On remarque, pour la classe I, une amélioration intéressante de la précision. Mais les progressions de précision les plus notables et les plus significatives sont celles des classes E et G. En essayant d'expliquer ce constat, on est rentré plus dans le détail de calcul des distances pour chaque fonction de Morse sur E et G. On a découvert que sans l'apprentissage et la combinaison, la fonction distance au centre de masse n'était pas du tout pertinente sur le type d'images des classes E. Ceci faussait complètement la distance globale lorsqu'on sommait les différentes distances sans leur donner de poids. Même constat pour la classe G mais cette fois-ci c'était la fonction cosinus de l'angle formé entre le vecteur tangent à la forme et l'axe majeur, qui n'était pas pertinente pour le type d'images dans G. Après l'apprentissage et la pondération des distances, selon notre méthode de calcul des poids des

comparaisons intra-classe, le poids de la distance des fonctions non-pertinentes est devenu très faibles d'où une meilleure précision de comparaison intra-classe. On a, à partir de ça, compris que si l'amélioration de la précision dans les autres classes n'était pas spectaculaire, c'était du au fait que le choix des fonctions de Morse était assez judicieux ou que leurs pertinences respectives se valaient à certain égards.

3.8 Conclusion

Dans ce chapitre une définition formelle du descripteur de forme basé sur la théorie de Morse a été présentée. Plusieurs théorèmes et propositions qui démontrent la puissance de ce descripteur y sont expliqués. Notamment, une étude très intéressante concernant l'approximation de n'importe qu'elle fonction réelle et continue par une fonction de Morse. Ce résultat pourrait étendre la définition du notre descripteur à une infinité d'autre fonctions ce qui pourrait très bien, constituer une base à d'autres descripteurs. Les résultats des expériences que nous avons menés, même comparés à des standards dans le domaine, ont prouvé la validité de l'utilisation du descripteur de forme basé sur la théorie de Morse.

Conclusion

Nous avons effectué dans ce mémoire une étude du domaine de la description et l'analyse de forme. Une attention plus accrue a été portée aux propriétés topologiques d'une forme. Car ces mêmes propriétés sont la base même de notre descripteur.

Un travail de synthèse et de classification des descripteurs de forme déjà existant a, tout d'abord, été effectué, et ce, pour clarifier les différentes approches utilisées et les différentes propriétés d'une forme extraites de l'image, pour permettre de mieux la caractériser. Ce travail de synthèse nous a permis de motiver l'utilisation de nouvelles propriétés de l'image. Celles-ci sont bien évidemment les propriétés topologiques telles que l'homotopie et l'homologie. Nous avons, par la suite, introduit les notions importantes de la topologie algébrique et de la théorie de Morse. Nous avons choisi de présenter les résultats les plus utiles et pratiques pour la description d'une forme.

Le lien entre fonctions de Morse définies sur une forme, leurs points critiques notamment, et l'évolution de l'homologie de cette même forme, ce lien, est à la base du *descripteur de forme basé sur la théorie de Morse*. Nous avons démontré, théoriquement, à travers les propriétés combinées de la théorie de Morse et de la topologie computationnelle, la puissance d'un tel descripteur construit sur une telle base. Mais notre démonstration ne s'est pas arrêtée à la théorie, car un outil de recherche d'images par le contenu (CBIR) a été construit à base de ce descripteur. Des expérimentations standard à base de Précision/Rappel ont été menées sur une base de données de forme de poissons. En variant les critères de pertinences de résultats et en prenant notamment en compte l'aspect perceptuel, les résultats obtenus sont très encourageants et démontrent la pertinence de l'utilisation des propriétés topologiques pour définir une forme.

Malgré une approche heuristique satisfaisante, nous ne fournissons pas de moyen théorique lors du choix des fonctions de Morse pour définir une forme. En effet, ce choix est primordial pour construire un bon descripteur de forme. Ce choix dépend évidemment de la nature de la forme à décrire mais aussi de celles des autres formes avec lesquelles on comparera notre forme. Un travail théorique pourrait être fait dans ce domaine.

Parmi les résultats intéressants que nous avons présentés dans ce mémoire, figure celui qui décrit comment approximer n'importe quelle fonction réelle et continue par une fonction de Morse. Ce résultat pourrait servir à construire un autre descripteur qui offrirait un choix plus vaste de fonctions de mesure. Nous proposons aussi, pour optimiser notre descripteur, d'explorer les pistes décrites dans notre mémoire pour prendre en compte la persistance des points critiques et notamment la persistance homologique.

Ce domaine des descripteurs de forme reste, comme on le voit, vaste et propice à de plus amples explorations. Notamment à cause de l'augmentation constante du volume de données mais aussi aux avancées dans la puissance des calculateurs qui nous permettent de faire des tentatives plus audacieuses.

Bibliographie

- [1] Course by Bryan S. Morse, Brigham Young University. Lecture 9: Shape Description (Regions).
http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/region-props-and-moments.pdf. In CVonline: On-Line Compendium of Computer Vision. R. Fisher(ed). Disponible: "<http://homepages.inf.ed.ac.uk/rbf/CVonline/>". [Date d'accès Février 2010]
- [2] S. Loncaric. *A Survey of Shape Analysis Techniques*. Pattern Recognition, Vol. 31, No. 8, 1998, pp. 983-1001,
- [3] T. Pavlidis, *A Review of Algorithms for Shape Analysis*, Computer Graphics and Image Processing, 7 (1978), pp. 243-258.
- [4] R. Duda and P. Hart. *Descriptions of Line and Shape*, chapter 9, pp 327-378. John Wiley, 1973.
- [5] R. Nevatia. *Shape Analysis and Recognition*, chapter 5, pp 61-89. Prentice Hall, 1982.
- [6] L. Davis. *Two-dimensional shape representation*. In I. Young and K. S. Fu, editors, Handbook of Pattern Recognition and Image Processing, pages 233-245. Academic Press, 1986.
- [7] C. Zahn and R. Roskies. *Fourier descriptors for plane closed curves*. Computer Graphics and Image Processing, 21:269-281, 1972.
- [8] E. Persoon and K. Fu. *Shape discrimination using fourier descriptors*. IEEE Transactions on SMC, 7 : 170-179, 1977.
- [9] H. Blum, *A transformation for extracting new descriptors of shape*, in: W. Whalen-Dunn (Ed.), Models for the Perception of Speech and Visual Forms, MIT Press, Cambridge, MA, 1967, pp. 362-380.

- [10] M. R. Teague. *Image analysis via the general theory of moments*. Journal of the Optical Society of America, 70 : 920-930, 1980.
- [11] R. J. Prokop and A. P. Reeves. *A survey of moment-based techniques for unoccluded object representation and recognition*. CVGIP: Graphical Models and Image Processing, 54:438-460, 1992.
- [12] H. Kim, K. Park, and M. Kim. *Shape decomposition by collinearity*. Pattern Recognition Letters, 6:335-340, 1987.
- [13] H. V. Jagadish and A. M. Bruckstein. *On sequential shape descriptions*. Pattern Recognition, 25:165-172, 1992.
- [14] YANG Mingqiang. *Shape-based feature extraction and similarity matching* Ph.D thesis, Ecole doctorale MATISSE, Rennes, Juillet 2008.
- [15] D. Zhang and G. Lu, *Review of Shape Representation and Description Techniques*, Pattern Recognition, 2004,37(1):1-19
- [16] I. Yong, J. Walker, J. Bowie, *An analysis technique for biological shape*, Comput. Graphics Image Process. 25 (1974) 357-370
- [17] M. Peura and J. Iivarinen, *Efficiency of simple shape descriptors*, in Proc. 3rd International Workshop on Visual Form (IWVF3), May 1997.
- [18] E.R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*, Academic Press, New York, 1997, pp. 171-191.
- [19] K.L. Kashyap, R. Chellappa, *Stochastic models for closed boundary analysis: representation and reconstruction*, IEEE Trans. Inform. Theory 27 (1981) 627-637.
- [20] R. Chellappa, R. Bagdazian, *Fourier coding of image boundaries*, IEEE Trans. Pattern Anal. Mach. Intell. 6 (1) (1984) 102-105.
- [21] D. Zhang and G. Lu, *A comparative study of fourier descriptors for shape representation and retrieval*, in Proc. 5th Asian Conference on Computer Vision, 2002.
- [22] D. Zhang and G. Lu, *A comparative study of curvature scale space and fourier descriptors for shape-based image retrieval*, Visual Communication and Image Representation, vol. 14(1), 2003.

- [23] T. Pavlidis, *Algorithms for Graphics and Image Processing*, Computer Science Press, Rockville, MD, 1982, p. 143.
- [24] H. Freeman, *On the encoding of arbitrary geometric configurations*, IRE Trans. Electron. Comput. EC-10 (1961) 260–268.
- [25] W.I. Groskey, P. Neo, R. Mehrotra, *A pictorial index mechanism for model-based matching*, Data Knowledge Eng. 8 (1992) 309–327.
- [26] R. Mehrotra, J.E. Gary, *Similar-shape retrieval in shape data management*, IEEE Comput. 28 (9) (1995) 57–62.
- [27] S. Berretti, A.D. Bimbo, P. Pala, *Retrieval by shape similarity with perceptual distance and effective indexing*, IEEE Trans. Multimedia 2 (4) (2000) 225–239.
- [28] M. E. Celebi and Y. A. Aslandogan, *A comparative study of three moment-based shape descriptors* in Proc. of the International Conference of Information Technology: Coding and Computing, 2005, pp. 788-793.
- [29] D.S. Zhang, G. Lu, *Generic Fourier descriptor for shape-based image retrieval*, in: Proceedings of IEEE International Conference on Multimedia and Expo (ICME2002), Vol. 1, Lausanne, Switzerland, August 26–29, 2002, pp. 425–428.
- [30] A. Goshtasby, *Description and discrimination of planar shapes using shape matrices*, IEEE Trans. Pattern Anal. Mach. Intell. 7 (1985) 738–743.
- [31] A. Taza, C. Suen, *Discrimination of planar shapes using shape matrices*, IEEE Trans. Syst. Man Cybern. 19 (1989) 1281–1289.
- [32] R. Datta, D. Joshi, J. Li, and J. Z. Wang, *Image retrieval: Ideas, influences, and trends of the new age*, Penn State University, Tech. Rep., 2006.
- [33] R. C. Veltkamp and M. Hagedoorn, *State-of-the-art in shape matching*, Utrecht University, Tech. Rep., 1999.
- [34] J. Ramon, *Image analysis and recognition*. Springer, 2004, ch. Finding significant points for a handwritten classification task, pp. 440-460.
- [35] B. Scassellati, S. Alexopoulos, and M. Flickner, *Retrieving images by-2D shape: a comparison of computation methods with human perceptual judgments*, in Proc. SPIE Storage and Retrieval for Image and Video databases, 1994.

- [36] S. Krishnamachari and M. A. Mottaleb, *Compact color descriptor for fast image and video segment retrieval* in Proc. SPIE Storage and Retrieval for Media Databases, 2000.
- [37] T. Kaczynski, K. Mischaikow, M. Mrozek : Computational Homology. Appl. Math. Sci. Series 157, Springer-Verlag, New York 2004.
- [38] P. J. Giblin. *Graphs, Surfaces and Homology*. Chapman and Hall, London, 1977.
- [39] William S. Massey. *A Basic Course in Algebraic Topology*. Springer-Verlag, New York, 1980.
- [40] M.J. Greenberg and J.R. Harper. *Algebraic Topology: A First Course*. Benjamin-Cummings, 1981.
- [41] J. R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, Cambridge-Massachusetts, 1984.
- [42] Edwin H. Spanier. *Algebraic Topology*. Springer-Verlag, New York, 1994.
- [43] A. Hatcher. *Algebraic Topology*. Cambridge University Press, New York, 2001.
- [44] S. Derivière, A. Trahan et T. Kaczynski. Comment compter les trous dans une meule de fromage suisse : ou, l'homologie pour les gourmands. *Congrès AMQ2006, Sherbrooke*. 2006.
- [45] Corriveau, David, "*Méthodes topologiques pour l'étude des formes*" Thèse (M.Sc.) Université de Sherbrooke. 2004.
- [46] Kaczyński, T., Mrozek, M., Ślusarek, M., *Homology computation by reduction of chain complexes*. Computers and Math. Appl. 1998, Pages 59-70
- [47] M. Allili, D. Ziou, *Computational Homology Approach for Topology Descriptors and Shape Representation*, Proceedings of International Conference on Image and Signal Processing (ICISP'2003), vol. 2, 2003, pp. 508-516.
- [48] M. Allili, K. Mischaikow, and A. Tannenbaum. *Cubical homology and the topological classification of 2d and 3d imagery*. Proceedings of ICIP 2001, 2:173–176, 2001.
- [49] T. Kaczynski, K. Mischaikow, and M. Mrozek. *Applied Mathematical Sciences: Computational Homology*. Springer Verlag, New York, 2004.

- [50] D. Ziou and M. Allili. *Generating Cubical Complexes from Image Data and Computation of the Euler Number*. Pattern Recognition, 35:2833–2839, 2002.
- [51] Milnor, J. Morse Theory, vol. 51 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1963.
- [52] Bott, R. *Morse theory and its application to homotopy theory*. Universitat Bonn, Germany, 1960.
- [53] M. Allili, D. Corriveau, *Topological analysis of shapes using Morse theory*, CVIU 105:188–199, 2007.
- [54] F. Cagliari, C. Landi, and L. Grasselli. Presentations of morse homology for studying shapes of manifolds. *preprint*, 2001.
- [55] Patrizio Frosini. *Connections between size functions and critical points*. Mathematical Methods in the Applied Sciences, 19:555–569, 1996.
- [56] A. Verri and C. Uras. *Metric-topological approach to shape representation and recognition*. Image and Vision Computing, 14:186–207, 1996.
- [57] Y. Matsumoto. *An Introduction to Morse Theory*, volume 208 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2002.
- [58] H. Edelsbrunner and J. Harer, *Persistent homology - a survey.*, In Twenty Years After, eds. J. E. Goodman, J. Pach and R. Pollack, AMS. (2007)
- [59] S. Abbasi, F. Mokhtarian, and J. Kittler, *SQUID Demo Dataset 1,500*, <http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html>, 1997.
- [60] F. Mokhtarian and A.K. Mackworth, *A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 14, no. 8, pp. 789-805, Aug. 1992.
- [61] <http://www.ee.surrey.ac.uk/CVSSP/demos/css/demo.html> (Avril 2011).
- [62] L.J. Latecki, R. Lakaemper, and U. Eckhardt, *Shape Descriptors for Non-Rigid Shapes with a Single Closed Contour*, Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '00), pp. 1424-1429, June 2000.
- [63] Milios E., Petrakis E.: *Shape Retrieval Based on Dynamic Programming*, IEEE Transactions on Image Processing, January 2000, Number 1, Volume 1, pp. 141-147.

- [64] C. Uras and A. Verri. *Studying Shape Through Size Functions*. Shape in Picture, NATO ASI Series F 126, pages 81–90, 1994.
- [65] F. Cagliari. *Presentations of Morse Homology for Studying Shape of Manifolds*. preprint, 2001.